Biologically Active Peptides from Australian Amphibians

A thesis submitted for the Degree of Doctor of Philosophy

by

Rebecca Jo Jackway B. Sc. (Biomed.) (Hons.)

from the

Department of Chemistry, The University of Adelaide

August, 2008

Chapter 1 Amphibian Bioactive Peptides

1.1 Anuran Skin Secretions

For centuries, the skin extracts of anurans (frogs and toads) have been used in folk medicine and witchcraft. The very first reference of use was by the Assyrian and Babylonian priests over three to four thousand years ago [1]. Since then, frog extracts have been used commonly in Chinese folk medicine to regulate internal bodily functions and fertility [1] and Ancient Egyptians used extracts of *Paparer somniferum* to treat pain and diarrhoea [2]. Amazonian Indians also used frog secretions in hunting to increase awareness and also as poisons on their hunting tools [3, 4]. There are endless references to the use of frog extracts to produce euphoria and hallucinations, particularly in religious ceremonies [1, 2]. In addition, Shakespeare referred to the used of toad venoms in witches' brews to alter human behaviour in Macbeth [2]. With the development of modern science, investigation into the structure and pharmacological function of the active component present in the skin extracts has become vital in order to develop new pharmaceuticals.

Amphibians are believed to have developed from freshwater fish at the end of the Devonian period [1]. The transformation from an aquatic to a terrestrial environment required numerous morphofunctional and behavioural adaptations to ensure the survival of amphibians in their new environment [5]. An important adaptation was the development of glands on the epidermis that play an integral role in regulation of the skin physiology and primary host defence.

The rapid global decline in amphibian populations due to habitat destruction, introduction of new predators, global warming and pollution, is a growing concern [6, 7]. In addition, the worldwide spread of foreign pathogens such as the chytrid fungus *Batrachochytrium dendrobatidis*, which is native in amphibians in Australia and the Americas, has resulted in large population declines and species extinction in Europe and Africa, in addition to Central America and Australia [8, 9]. Since the skin is the first barrier against such pathogen invasion, the secretory products released by the dorsal granular glands constitute an essential frontline defence mechanism.

Amphibian skin secretions include a rich chemical arsenal of diverse components including biogenic amines, peptides, alkaloids and bufadienolides [10, 11], which display antimicrobial, hormone, analgesic and neurotransmitter actions [12]. In addition to these compounds, several enzymes involved in the processing of inactive propeptides into their corresponding bioactive peptides have been isolated [5][1*.]. The concentration of these components is highest in the skin secretions, however many have also been found in the blood, internal organs and in the eggs and tadpoles of some amphibian species [5, 13]. The peptides and other noxious components in the skin secretion have significantly contributed to the adaption of amphibians within hostile environments.

In the 1960's Erspamer pioneered investigation into characterisation of anuran skin secretions and established that the secretion was an extraordinary source of biologically active peptides. Studies of some 500 distinct species have illustrated unexpected diversity in the identified peptides, with different species secreting a different spectrum of bioactive peptides. Among these active peptides are neuropeptides, peptides with antimicrobial, antiviral and fungicide activity, peptides that complex with the regulatory protein calmodulin (CaM) and sex pheromones [14]. To date, no two species have been identified as having identical skin secretion profiles. Nevertheless, particular peptides have been shown to be prevalent in different species residing in certain geographical regions. For example, bombesin-like peptides are found in numerous species in Europe and Asia, whilst caerulein peptides are found in species distributed throughout Central and South America, Australia and southeastern Africa [15]. At present, only a fraction of the world's amphibian species have been investigated, thus still leaving an abundance still be to studied to reveal new active components with valuable biological properties.

Erspamer demonstrated that amphibian skin secretions contain a variety of peptides that are homologous to mammalian hormones and neurotransmitters. However, the quantities of these peptides from skin secretions are generally greater than that found in any mammalian tissue [15, 16]. For example, caerulein isolated from numerous *Litoria* species is an analogue of cholecystokinin (CCK) and gastrin, which are expressed in the brain and gastrointestinal (GI) tract of mammals [10]. Similarly, physalaemin [17], phyllomedusin [18] and angiotensin [10] are analogues of substance P, whilst xenopsin [19] is analogous

[1*.] Clark, D.P., Durell, S., Maloyi, W.L. and Zasloff, M. (1994) Ranalexin: A novel antimicrobial peptide from bullfrog (*Rana catesbeiana*) skin, structurally related to the bacterial antibiotic, polymyxin, *J. Biol. Chem.* 269, 10849-10855.

to neurotensin [15]. Other peptides such as thyroptropin-releasing hormone (TRH) and bombesin peptides are found in both amphibians and mammals [16].

A number of amphibian peptides have led to the development of pharmaceutical agents that are used today. These have formed treatments for numerous medical conditions such as GI infections, foot ulcers, conjunctivitis and cancers [20]. Epibatidine, an alkaloid isolated from Ecuadorian frogs, formed the lead compound for the development of non-opioid analgesics that lack side effects such as respiratory depression, constipation and physical dependence that are normally associated with opioid analgesics [21]. Caerulein has been used for decades to treat a variety of GI conditions and as part of a diagnostic technique to identify GI problems [22]. Furthermore, with the increase in antibiotic resistance making treatment of infection difficult, amphibian peptides present promise as antibiotics that will be void of any bacterial resistance. Since the discovery of magainin peptides from *Xenopus laevis* [23], the main research interests have been focused on antimicrobial peptides. Analogues of magainin are currently being developed as broad spectrum topical antibiotic agents and are within clinical trials, soon to be on the market as human drugs for cutaneous infections [24].

1.2 Peptides from Australian Frogs

A major focus of the Bowie research group for the last two decades has been identification and characterisation of the biologically active peptides in the skin secretions of frogs from Australasian genera *Crinia*, *Cyclorana*, *Litoria*, *Limnodynastes* and *Uperoleia*. There are several thousand anuran species worldwide, but only some 220 in Australasia. Many of these species are indigenous to Australia, some having evolved following or during the period from 100 to 80 million years ago when Australia separated from Gondwanaland [25]. In 1984 Erspamer and coworkers completed an extensive study of one hundred Australasian amphibian species. The extract preparations isolated from the dried skins were subjected to biological screening on smooth muscle preparations and for effects on systemic blood pressure [26]. Erspamer identified the caerulein, bombesin and tachykinin neuropeptides as the most frequent and abundant peptides.

The Bowie research group has isolated and identified the biologically active peptides from the dorsal glands of more than 35 species of Australian amphibians. These peptides have

been isolated from metamorphic and adult animals. At least one tadpole species (*Litoria splendida*) displays the same skin peptide profile as the adult animal [13]. Many Australian frogs contain a number of major components in their skin peptide profiles including a smooth muscle active agent that sometimes also possesses analgesic activity, a mixture of narrow and broad spectrum antimicrobial peptides, at least one neuropeptide and a neuronal nitric oxide synthase (nNOS) inhibiting peptide. Peptides possessing antiviral, anticancer and fungicide activities, hormones and a sex pheromone have also been identified [14, 27]. Many peptides show multifaceted activity, whilst the activity of others is still unknown. Table 1.1 illustrates the sequences and activities of selected peptides isolated from Australian frogs.

Neuropeptides and peptides that complex with CaM to inhibit nNOS formation of nitric oxide (NO) will not be discussed here. These will be discussed in detail in Chapters 5 and 6. In this section, only antimicrobial peptides and pheromones isolated from Australian amphibian species will be discussed.

Table 1.1: Selected peptides isolated from Australian amphibian skin secretions. Adapted from [27].

Peptide	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
Aurein 2.2	GLFDIVKKVVGALGSL-NH_2	b	1-3
Aurein 3.1	GLFDIVKKIAGHIAGSI-NH_2	a, b	1, 2
Aurein 4.1	GLIQTIKEKLKELAGGLVTGIQS-OH	b	1, 2
Aurein 5.2	GLMSSIGKALGGLIVDVLKPKTPAS-OH	a, b	1, 2
Caeridin 1.1	GLLαDGLLGTGL-NH_2	c, d, e, f, g	
Caeridin 1.2	GLLβDGLLGTGL-NH_2	e	
Caeridin 2	GLLDVVGNLLGGLGL-NH_2	d, e	
Caeridin 3	GLFDAIGNLLGGLGL-NH_2	d, e	
Caeridin 4	GLLDVVGNVLHSGL-NH_2	d	
Caerin 1.1	GLLSVLGSVAKHVLPHVVPVIAEHL-NH_2	c, d, e	1-5
Caerin 1.3	GLLSVLGSVAQHVLPHVVPVIAEHL-NH_2	d	1, 2
Caerin 1.4	GLLSSLGSVAKHVLPHVVPVIAEHL-NH_2	d, e	1
Caerin 1.8	GLFKVLGSVAKHLLPHVVPVIAEKL-NH_2	g	1-3, 5
Caerin 1.9	GLFGVLGSIAKHVLPHVVPVIAEKL-NH_2	g	1-5
Caerin 2.1	GLVSSIGRALGGLLADVVKSKGQPA-OH	c	1, 5
Caerin 2.2	GLVSSIGRALGGLLADVVKSKEQPA-OH	d	1, 5
Caerin 3.1	GLWQKIKDKASELVSGIVEQVK-NH_2	c, d	1
Caerin 3.2	GLWEKIKEKASELVSGIVEGVK-NH_2	d	1
Caerin 4.1	GLWQKIKSAAGDLASGIVEGIKS-NH_2	d	1
Caerin 4.2	GLWQKIKSAAGDLASGIVEAIKS-NH_2	d	1
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	1-3, 5
Citropin 1.2	GLFDIIKKVASVVGGL-NH_2	i	1-3, 5
Citropin 1.3	GLFDIIKKVASVIGGL-NH_2	i	1-3, 5
Dahlein 1.1	GLFDIIKNIVSTL-NH_2	j	1
Dahlein 1.2	GLFDIIKNIFSGL-NH_2	j	1
Dahlein 4.1	GLWQLIKDKIKDAATGLVTGIQS-NH_2	j	
Dahlein 5.1	GLLGSIGNAIGAFIANKLKP-OH	j	5
Dynastin 1	GLLSGLGL-OH	k	
Dynastin 2	GLLSSLGLNL-OH	l	
Dynastin 3	GLVPNLLNNLGL-OH	m	
Dynastin 4	GLVSNLGI-OH	n	
Electrin 2.1	NEEEKVKWEPDVP-NH_2	o	
Fletcherin	AGPVSKLVSGIGL-OH	p	
Frenatin 1	GLLDALSGILGL-NH_2	q	
Frenatin 2	GLLGTGNLLNGLGL-NH_2	q	
Frenatin 3	GLMSVLGHAVGNVLGGLFKPKS-OH	q	5

Table 1.1: continues.

Peptide	Sequence	Species[1]	Activity[2]
Lesueurin	GLLDILKKVGKVA-NH_2	r	
Maculatin 1.1	GLFGVLAKVAAHVVPAIAEHF-NH_2	s	1-3, 5
Maculatin 1.3	GLLGLLGSVVSHVVPAIVGHF-NH_2	t	1, 2
Maculatin 2.1	GFVDFLKKVAGTIANVVT-NH_2	t	1, 2
Maculatin 3.1	GLLQTIKEKLESLAKGIVSGIQA-NH_2	s	
Riparin 1.1	RLCIPVIFPC-OH	u	6
Riparin 2.1	IIEKLVNTALGLLSGL-NH_2	u	1
Rothein 2.1	AGGLDDLLEPVLNSADNLVHGL-NH_2	v	
Rothein 3.1	ASAAGAVRAGGLDDLLEPVLNSADNLVHGL-NH_2	v	
Rubellidin 4.1	GLGDILGLLGL-NH_2	w	
Rubellidin 4.2	AGLLDILGL-NH_2	w	
Signiferin 1	RLCIPYIIPC-OH	x	6
Signiferin 2.1	IIGHLIKTALGMLGL-NH_2	x	1
Signiferin 3.1	GIAEFLNYIKSKA-NH_2	u, x	5
Splendipherin	GLVSSIGKALGGLLADVVKSKGQPA-OH	c, d	7
Tryptophyllin L 1.1	PWL-NH_2	w	
Tryptophyllin L 1.2	FPWL-NH_2	o, w	
Tryptophyllin L 1.3	pEFPWL-NH_2	w	6
Tryptophyllin L 1.4	FPFPWL-NH_2	w	
Uperin 1.1	pEADPNAFYGLM-NH_2	y	6
Uperolein	pEPDPNAFYGLM-NH_2	z	6

[1]**Species: (a)** *Litoria raniformis* [28]; **(b)** *Litoria aurea* [28]; **(c)** *Litoria splendida* [29]; **(d)** *Litoria caerulea* [30]; **(e)** *Litoria gilleni* [31]; **(f)** *Litoria xanthomera* [32]; **(g)** *Litoria chloris* [33]; **(h)** various species from genus *Litoria*, *Xenopus laevis* [10, 34], *Leptodactylus labyrinthicus*; **(i)** *Litoria citropa* [35]; **(j)** *Litoria dahlii* [36]; **(k)** *Limnodynastes interioris* [37]; **(l)** *Limnodynastes dumerilii* [37]; **(m)** *Limnodynastes terraereginae* [37]; **(n)** *Limnodynastes salmini* [38]; **(o)** *Litoria electrica* [39]; **(p)** *Limnodynastes fletcheri* [38]; **(q)** *Litoria infrafrenata* [40], **(r)** *Litoria lesueurri* [41]; **(s)** *Litoria genimaculata* [42]; **(t)** *Litoria eucnemis* [43]; **(u)** *Crinia riparia* [44]; **(v)** *Litoria rothii* [45]; **(w)** *Litoria rubella* [46]; **(x)** *Crinia signifera* [47]; **(y)** *Uperoleia inundata* [48]; **(z)** many species of the genus *Uperoleia* [10].

[2]**Activity: (1)** antimicrobial agent; **(2)** anticancer agent; **(3)** fungicide; **(4)** antiviral; **(5)** nNOS inhibitor; **(6)** neuropeptide (displaying activities such as smooth muscle contraction, immunomodulation, hormone, neurotransmitter and opioid activity) (See Chapter 6); **(7)** pheromone.

1.2.1 Antimicrobial Peptides

Typically, frogs live in environments that are rich in microbial flora, against which they require an adequate defence immune mechanism [49]. Antimicrobial peptides form an important frontline defence against invading microbes [50]. Generally, amphibian species produce a unique collection of antimicrobial peptides that display similar structural features despite targeting specific microorganisms [51]. It is believed that anurans produce the range of antimicrobial peptides that work in synergy in order to: (i) provide a better shield against a wide spectrum of microbes including bacteria, fungi and protozoa; and (ii) to minimise the formation of any resistant microbial strains. So far, no two amphibians of different families, genera or species produce precisely the same spectrum of antimicrobial peptides. Amphibians produce antimicrobial peptides that show diversity in length, hydrophobicity, charge and spectrum of activity.

Amphibians produce around 20 % of the known antimicrobial peptides. Typically, these peptides are 11 to 46 residues long [52], positively charged due to the high content of Lys and Arg residues and are encoded by genes displaying high homology between species [53]. These antimicrobial peptides are generally classified into two broad classes based on their chemical structure; linear α helical peptides or cyclic peptides containing C-terminal disulfide links [53, 54]. A common characteristic of antimicrobial peptides is the propensity to form amphiphathic α helices. The positively charged regions of the helix bind to the negatively charged lipid bilayers of bacterial membranes, thereby disrupting them. Interestingly, antimicrobial peptides display activity against a broad spectrum of microorganisms including bacteria, fungi, viruses and malaria parasites [52], in addition to antitumour activity [14].

Many of the antimicrobial peptides isolated from amphibians are C-terminally amidated. An exception to this however, is magainin, which carries a free acid [23]. It has been demonstrated that peptides that have C-terminal amides generally display greater antimicrobial activity [50]. An increased positive charge usually results in an increase in antimicrobial activity, however a large positive charge (eg. greater than +9) can reduce potency [55]. This reduction in activity may be a consequence of the higher charge density affecting the propensity of helix formation [56].

It has been demonstrated that substitution of the residues in the native antimicrobial peptide with their corresponding D-isomers resulted in biological activities that were indistinguishable [11, 27], thus these peptides operate in a receptor-independent mode. The mechanism of action of antimicrobial peptides occurs in two main steps, the first being association with the cellular surface and second, membrane permeabilization. The cationic antimicrobial peptides are largely unstructured in aqueous solutions, however on electrostatic interaction with the anionic components of the target pathogen membrane, the peptides form amphipathic helices [54, 57]. The peptides insert into the membrane and disturb its function in two main ways; carpet [58] and the barrel-stave mechanisms [59].

The carpet model describes the binding of the positively charged regions of the helices to the negatively charged phospholipids in the membrane. The peptides assimilate into the membrane and cause an increase in its surface area. This strains the bilayer and results in the formation of transient pores. The membrane finally collapses into small vesicles coated in the peptide (Figure 1.1b) [60]. Alternatively, the barrel-stave model describes the aggregation of helical peptides on the membrane surface [61]. The peptides assume an orientation that allows their hydrophobic regions to interact with the hydrophobic phospholipid core. Consequently, the hydrophilic peptide surfaces orient inwards and form aqueous channels that disrupt the osmotic potential of the cell (Figure 1.1a) [59, 62]. For this model, a minimum of 20 residues is required to span the membrane, although for smaller peptides a modified model has been proposed whereby these peptides dimerise end on to effect complete penetration [63]. In both models, the disruption of the membrane leads to the breakdown of the transmembrane potential and ion gradients, creating leakage of the cellular contents resulting in cell death.

Figure 1.1: (a) Barrel-stave model; the helical peptides aggregate on the membrane surface prior to forming aqueous channels, and **(b)** carpet model; the peptides assimilate into the membrane and the membrane dissociates when a threshold concentration is reached. The helical peptides are represented by cylinders with hydrophobic and hydrophilic regions represented by red and blue respectively. Adapted from [64].

Bacterial resistance to antibiotics has become a major concern as the number of resistant strains increases at a faster rate than the introduction of new antibiotic agents [65]. Invading microoganisms are not expected to develop resistance against the antimicrobial peptides as the cellular integrity of the microorganism is destroyed within a matter of minutes [50]. In addition, these peptides may not only target the cellular membrane integrity but also affect their antimicrobial activity by acting as an intracellular signal to coordinate the innate and adaptive host defence to initiate processes such as wound healing, angiogenesis and cell proliferation [66]. Thus, amphibian antimicrobial peptides provide attractive therapeutic prospects in modern medicine.

Australian amphibians have proved to be an abundant source of antimicrobial peptides including the aureins [28], citropins 1 [35], caerins [30, 31], dahleins 1 [36], frenatins [40] and maculatins [42, 43] (Table 1.1), many of which are small cationic peptides that adopt amphipathic α helical structures at the lipid bilayer of bacterial cells. In addition to their antimicrobial activities, many of these peptide types show fungicide activity against chytrid fungus (*Batrachochytrium dendrobatidis*) which is widespread among Australian anuran species [14]. The caerin peptides have C-terminal amides with their basic residues concentrated in the C-terminal region. Generally, these peptides are active only against Gram-positive bacteria [14]. Caerin 1.1 adopts a helix-hinge-helix conformer (Figure 1.2) [67] and it has been demonstrated that this peptide disrupts the bacterial membrane via the carpet mechanism [68, 69]. Caerin 1.1 is a wide spectrum antimicrobial agent against Gram-positive bacteria, however it has also been shown to be an anticancer agent against most types of human tumours at micromolar concentrations and active against enveloped viruses including HIV and *Herpes simplex 1* [27] at much lower concentrations than that which is toxic to cells [52]. It also is an antifungal agent, kills nematodes and inhibits the formation of NO by nNOS [14].

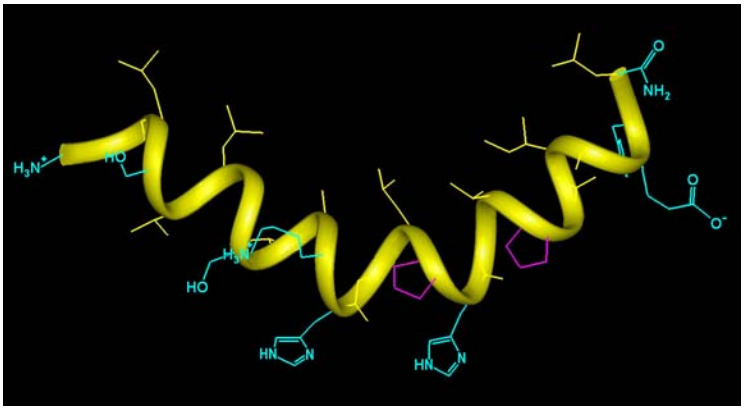


Figure 1.2: Solution structure of caerin 1.1 in TFE/H_2O. The hydrophobic residues are illustrated in yellow, whilst the hydrophilic residues are shown in blue. Pro residues are shown in pink [67].

1.2.2 Amphibian Pheromones

Numerous animals use chemical signalling (pheromones) to communicate [70-72]. Information received from pheromones released from a conspecific can alter the reproductive behaviour or physiology of an individual. Pheromones secreted from various glands are important in reproductive behaviour for sex recognition and courtship. The

effect of pheromones on sexual behaviour has been widely noted for a number of vertebrate species, however, very few have actually been isolated and identified [73].

In anurans, acoustical communication plays an integral role in sex recognition and mating [73]. However, because anurans often breed in aquatic environments and deposit their eggs in water, it is probable that anurans produce peptides as hydrophilic aquatic pheromones. This hypothesis was supported by evidence of aquatic sex pheromones in fish [74] and the isolation of aquatic sex pheromones from amphibians [75, 76]. The first sex pheromone isolated in amphibians was an aquatic sex pheromone, a peptide named sodefrin from the aquatic Japanese fire-bellied newt *Cynops pyrrhogaster* [75]. The male sex pheromone of a similar newt species, silefrin was subsequently isolated from the sword-tailed newt *Cynops ensicauda* [76]. The male releases the pheromone in the water and distributes it by vigorously waving its tail to attract the female [73].

Sodefrin	SIPSKDALLK-OH
Silefrin	SILSKDAQLK-OH

The first anuran sex pheromone named splendipherin (Figure 1.3, Table 1.1) was isolated from the male *Litoria splendida* [29, 77]. The skin secretions of both the male and female *L. splendida* were collected and analysed monthly over a three year period. The 25 residue peptide splendipherin was isolated from the male during the breeding months. Behavioural testing demonstrated that splendipherin release allows the male animal to be located and identified by the female [29].

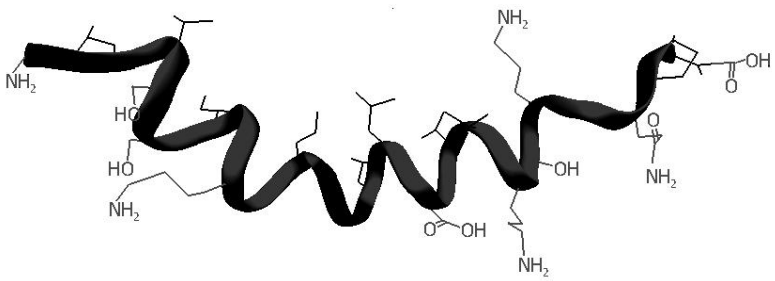


Figure 1.3: The solution structure of splendipherin in TFE/H_2O [78].

1.3 Collecting Skin Secretions

The skin of frogs contain two types of glands, namely mucous and granular glands (Figure 1.4). Mucous glands are scattered in large numbers across the entire surface of the skin and release a watery secretion containing salts such as Na^+ and Cl^- ions [79] that controls the physiology of the skin, including the maintenance of pH, respiration and moisture balance [5]. Alternatively, the granular glands are less abundant and predominately found on the dorsal skin surfaces. They release granules containing organic compounds and sterols [5, 79] and form the host's primary defence.

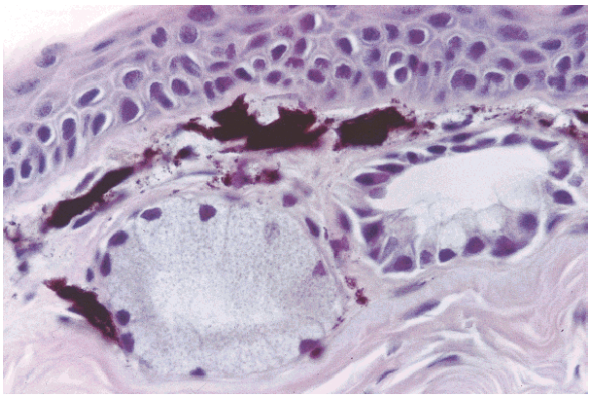


Figure 1.4: Light microscopic section of dorsal skin of a wild *Rana esculenta* stained with haematoxylin-eosin. The cross section shows a serous granular gland (left) and a mucous gland (right) [11].

Granular glands are comprised of densely packed secretory granules held in a syncytium, which is enclosed in a myoepithelial cell envelope. A duct formed by modified keratinocytes passes through the epidermis and allows the granular contents to be released onto the external surface of the animal (Figure 1.5) [5]. Contraction of the myoepithelial cell envelope compresses the syncytium and forces the granules, nuclei and associated organelles to be expelled on the external skin surface [80] by a holocrine secretory mechanism [1, 11, 81]. Sympathetic nerves control contraction of the myoepithelial cells, which enables the secretory contents to be expelled as a milky, viscous mucus in response to stress [5, 11]. Emptying of the granular glands is followed by a rearrangement of the secretory cells and resynthesis of the granular contents, occurring within days to weeks [10, 80, 81].

Figure 1.5: Schematic representation of a granular gland in *Xenopus laevis*. Adapted from [80].

Traditionally, the components of the skin secretions of anurans were isolated from the dried skin extractions of large numbers of animals. In some species, over 1,000 specimens were required to be sacrificed in order to obtain sufficient material for compound identification [82]. This method involved the skins to be removed, dried and extracted with organic solvents to acquire the contents of the skin glands [82]. Chemically, this procedure is inappropriate as many of the active peptides are stored in the corresponding inactive propeptides in the glands [14]. However, due to the alarming decline of amphibian numbers, non-harmful techniques are now required to obtain the skin secretions. One such technique involves the injection of either adrenaline or nonadrenaline directly into the dorsal sac [80]. Although this method is non-fatal, it is still relatively invasive. A more non-invasive method is the surface electrical stimulation technique, which requires gentle stroking of a platinum electrode over the dorsal surface and in turn applies a mild electrical stimulation to the dorsal glands [83]. An immediate discharge of the granular contents onto the skin surface occurs, which can be washed from the surface and collected (Figure 1.6). This technique is well tolerated by the animal and can be repeated on a regular basis every two to four weeks on a single animal with sufficient time for full replenishment of the glands. As a consequence, this technique is the preferred method used in our research.

Figure 1.6: Surface electrical stimulation technique for the collection of skin secretions of amphibians. The skin secretion is the white milky material at the tip of the electrodes.

Chapter 2 Methodology I - Mass Spectrometry

2.1 Introduction

Mass spectrometry (MS) is a technique used to measure accurate molecular mass and determine fragmentation patterns, which provide key information relating to the structure, composition and purity of a chemical compound. In order to do this, mass spectrometers must ionise the sample, separate the ions according to their mass-to-charge ratio (*m/z*) and then detect resulting ions [84].

The first mass spectrometer was developed by Sir J.J. Thomson in 1912 to illustrate the existence of stable isotopes [85]. This technique was based on the production, differentiation and detection of ions in the gas-phase. Mass spectrometers and their applications have advanced enormously since Thomson's fundamental experiments to create instruments with greater resolution and experimental capabilities. Time of flight (TOF) [86], quadrupole [87] and ion trap mass spectrometers [88] are now commonly used for the study of proteins and peptides.

All mass spectrometers use either electric or magnetic fields to ionise the sample. Until recently, MS was limited to relatively volatile species that could be ionised into the gas-phase via electron impact ionisation [89] or gas-phase chemical ionisation [90]. Recently, intact molecular ions for species of low molecular mass (3 – 5 kDa) have been ionised by fast atom bombardment (FAB) [91] and secondary ionisation mass spectrometry [92]. However, these techniques have limited biological application as they lack the sensitivity and the molecular mass range necessary to analyse large proteins. Major advances in biological mass spectrometric investigations have lead to the introduction of "soft" ionisation methods such as matrix-assisted laser desorption ionisation (MALDI) [93] and electrospray ionisation (ESI) (Section 2.3). These ionisation techniques produce $(MH)^+$ and/or $(M-H)^-$ species (together with multiply charged species) and allow direct analysis of thermally labile biological proteins and their complexes [94]. Compared to ESI, MALDI spectra are relatively poorly resolved and show lower sensitivity at high *m/z* values [94]. As a result, ESI has become widely adopted as a major method for analysing large, fragile and complex molecules including protein complexes and DNA fragments with mass in excess of several hundred kDa [95].

At present, MS is used widely in pharmaceutical, biotechnological, and bioanalytical research areas due to its high sensitivity, accuracy and rapid methods for structural characterisation of biological molecules [94]. This chapter describes the basic principles of MS and its applications to studying peptides and proteins in the gas-phase.

2.2 Q-TOF 2 Mass Spectrometer

The research presented in this thesis was carried out using a Q-TOF 2 hybrid quadrupole time of flight mass spectrometer. The principal components of the instrument can be seen in Figure 2.1. Q-TOF mass spectrometers combine the high performance of TOF analysis with the widely used technique of ESI to create high sensitivity and high mass accuracy. Originally targeted at the analysis of peptides, this instrument is now widely used for analysis of a variety of biological samples [96].

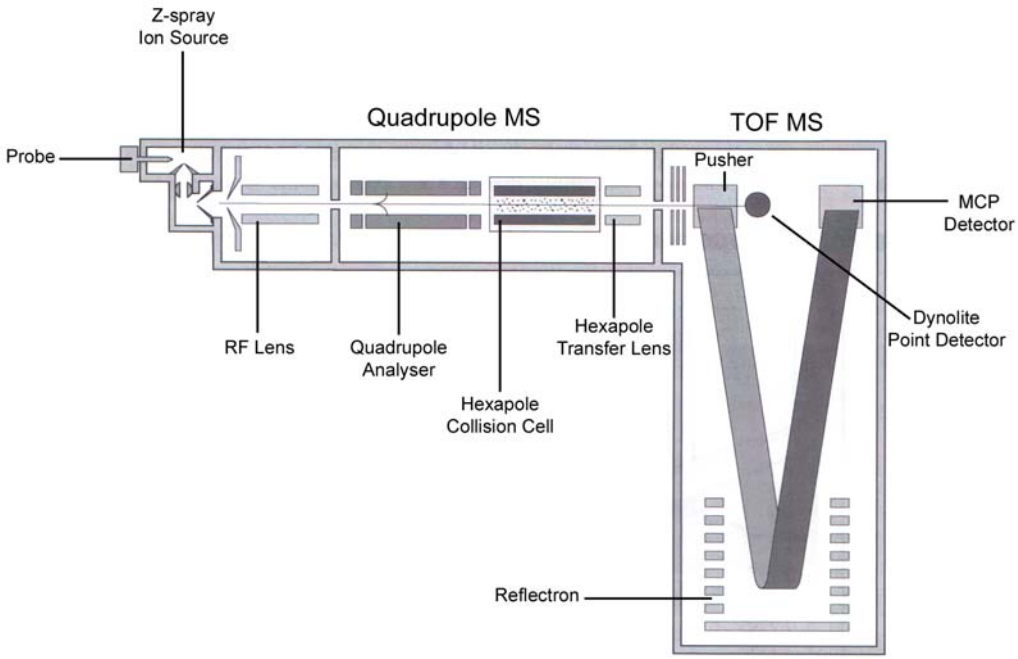


Figure 2.1: Schematic diagram of the Q-TOF 2 mass spectrometer [97].

Ionisation of gas-phase ions is achieved by ESI in the Z-spray ion source. The generated ions are pulled through the sample cone and into the spectrometer by application of an electric potential. Manipulation of the applied cone voltage allows for ions to be selected based on their *m/z*. Typically, larger *m/z* ions can be attracted through the cone into the

analyser when larger voltages are used [98]. This spectrometer consists of two sectors; the quadrupole and the TOF tube separated by a hexapole collision cell.

In the quadrupole sector the applied electric field is produced by four parallel rods, to which a fixed direct current (DC) and an alternating radiofrequency (*rf*) potential are applied. This potential difference focuses the ions into a beam that moves along the central axis of the rods. The motion of the ions through the quadrupole will depend on the electric field applied to the rods. Only ions with a particular *m/z* value will undergo stable motion and pass through to the detector. All other ions with unstable motion will be lost from the quadrupole (Figure 2.2) [87, 88, 99]. As a result, only a narrow *m/z* range is allowed to reach the detector [98]. By altering the strength and frequency of the applied DC and *rf* potentials, the applied electrical field can be varied to bring ions with different *m/z* values into focus on the detector and generate a mass spectrum [88, 99].

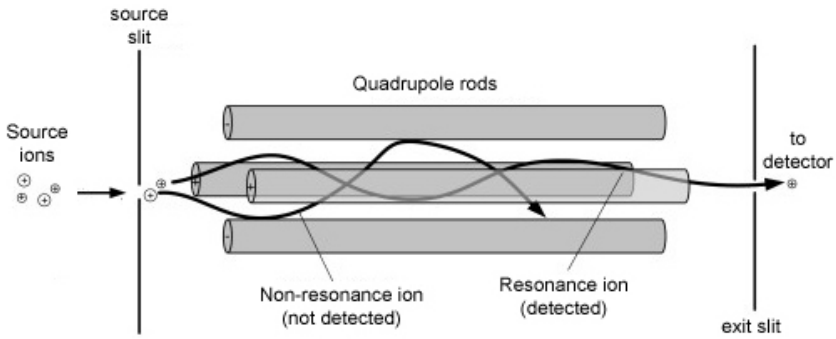


Figure 2.2: Schematic illustration of ion motion through the quadrupole sector.

In the MS mode, all of the ions are allowed to pass through the quadrupole mass analyser and accumulate at the pusher. The pusher accelerates the ions so that they have identical translational energy for entry into the TOF sector. The ions are accelerated down the TOF tube by an applied potential difference towards the reflectron, where they are electrostatically reflected back towards the detector. The ions are separated by the length of time that it takes the ion to reach the detector [100]. Translational energy is proportional to both mass and velocity squared, thus as the ions have the same translational energy, a heavier ion (larger *m/z*) will have a slower velocity and a longer TOF than a lighter ion [98]. If all the ions travel the same distance from the pusher to the detector, the difference in time can be interpreted into a mass spectrum [86]. Typically, TOF will achieve a resolution of 10,000 or greater (ie. able to separate *m/z* 1000.0 from 1000.1) and the

instrument has a *m/z* range to 4000 Th [98]. As spectra can be acquired from nanomolar concentrations, it is well suited for small quantities of peptide material obtained from natural sources.

If further structural information is required, tandem mass spectrometry (MS/MS) can be employed [101, 102]. In MS/MS mode, a particular ion of interest can be selected in the quadrupole analyser based on its *m/z* value and allowed to pass into the hexapole collision cell. The ion of interest is energised via a process known as collision induced dissociation (CID). The ions undergo collisions with noble gases such as Ar or He, which results in fragmentation of the parent ion at the most chemically labile bonds into a series of daughter ions [102-104]. The resultant daughter ions are then accelerated via the pusher into the TOF analyser for mass determination.

2.3 Electrospray Ionisation Mass Spectrometry

Until recently, using conventional ionisation methods, only low molecular weight, volatile compounds were amenable for analysis. The introduction of FAB-MS, then ESI-MS, refined the study of biological molecules as it allows the production of gas-phase ions of large and non-volatile molecules. In the case of ESI-MS this occurs directly from solution at atmospheric pressure [105]. The generation of ions by ESI was first established by Dole *et al*. in 1968 [106]. Later, these experiments were extended by Yamashita and Fenn and ESI-MS was introduced in 1984 [107, 108]. Although Dole's work included proteins as analytes and Fenn demonstrated that ESI could produce small multiply charged ions of small peptides, it was not until 1988 that ESI-MS was used routinely for the study of biological molecules [109].

In ESI-MS, the species of interest, initially in solution, can be ionised to form both positive and negative ions [110]. The solution is passed through a heated capillary at a low flow rate ($nL.min^{-1}$ to $\mu L.min^{-1}$). This not only reduces the amount of sample required, but also allows for the generation of smaller droplets, which results in an increase in signal level and more sensitive detection [111]. A voltage (2 - 3 kV) is applied to the metal needle (capillary) and an adjacent counter electrode causes a partial charge separation as the field penetrates below the surface of the liquid at the capillary tip [112]. Controlling the polarity bias of the applied field allows for positive or negative ions to be selected. For either

positive or negative ions, charge accumulation at the capillary tip causes the liquid surface to adopt a conical shape (Taylor cone) [112]. The tip of the Taylor cone is drawn into an elongated filament that becomes unstable when the charge density increases [105]. The electrical potential required for ESI is given by (1):

$$E_0 = \left(\frac{2\gamma \cos\theta}{\varepsilon_0 r_c} \right)^{1/2} \quad \textbf{(1)}$$

where γ is the liquid surface tension, θ = 49.3° and is the half-angle of the Taylor cone, ε_0 is the permittivity of the vacuum and r_c is the radius of the capillary [105, 112]. At this electrical potential, the filament breaks up into a mist of fine droplets (radius in micrometer range) [113], which travel down a potential and pressure gradient to the vacuum component of the mass spectrometer.

Solvent evaporation occurs as a result of the use of heating and dry gas application at a constant charge and causes a reduction in the radius of the droplet. The decrease in droplet radius R at a constant droplet charge q results in an increase in electrostatic repulsion of the charges at the droplet surface until the Rayleigh stability limit is reached [95, 105, 112].

$$q = 8\pi \left(\varepsilon_0 \gamma R^3 \right)^{1/2} \quad \textbf{(2)}$$

The Rayleigh equation (2) gives the conditions at which the electrostatic repulsion equals the force applied by the surface tension γ which is responsible for holding the droplet together. When the droplet radius and charge satisfy this equation, the charged droplet becomes unstable. After this point the coulombic repulsion overcomes the cohesive surface tension forces and the droplet divides [95, 105]. Continual solvent evaporation and coulombic fission results in the formation of very small droplets (radius in nanometer range) containing only a few elementary charges (Figure 2.3) [105, 112].

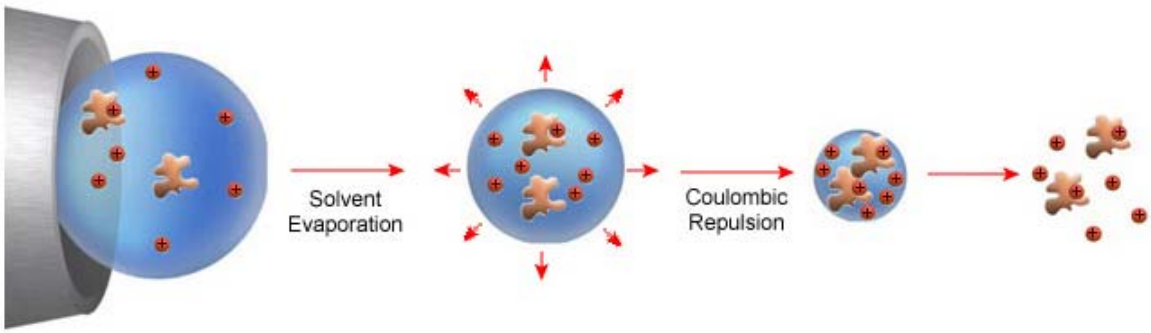


Figure 2.3: Schematic representation of the ESI process.

The exact mechanism involved in the formation of gas-phase ions from the small, highly charged droplets is unknown. Two mechanisms have been proposed; the charged-residue theory [106] and the ion evaporation model [114]. The charge-residue theory proposes that gas-phase ions are formed as a result of repeated coulomb fission forming smaller droplets until a point where no further solvent evaporation can occur and the droplets containing only one excess ion are formed [105]. In contrast, the ion evaporation model predicts that when the charged droplet radius has decreased to a region of 10 to 20 nm [110], direct ion emission will occur from the very small droplets due to electrostatic repulsion [105]. However, both models predict that the ions which are weakly interacting with the solvent will preferentially be expressed in the gas-phase and are most likely describing the mechanisms involved in formation of gas-phase ions in different situations [105].

2.4 Peptide Sequencing

MS has provided a method to determine the sequence of a peptide from minute amounts of sample. Traditionally, peptide sequencing was achieved through application of Edman sequencing (Section 2.4.4). However, this procedure can miss some residues. For example: (i) the last one or two residues may be missed; (ii) reactions at Pro can be slow; and (iii) some post-translational modifications [like pyroglutamate (pGlu)] are not detected. MS on the other hand, can be applied to identify these modifications. Furthermore, the sample does not need to be pure as the ion of interest can be isolated and independently analysed using MS/MS. Thus MS and MS/MS scans are particularly useful in ascertaining the sequence of peptide and proteins isolated from natural sources.

2.4.1 High Performance Liquid Chromatography

When a mixture of compounds is analysed by ESI-MS, the individual compounds compete for charge as they are produced from the droplets. As a result, only some of the most easily ionisable compounds will be observed [98]. Thus, there is a need to separate complex mixtures prior to MS analysis. The use of separation techniques such as high performance liquid chromatography (HPLC) preceding MS is widely used to analyse peptides from complex mixtures.

HPLC utilises the differences in properties of peptides, namely differences in size, shape, net charge and hydrophobicity. Generally, reverse phase HPLC is used and separates peptides due to differences in hydrophobicity. The peptides equilibrate between the hydrophobic stationary phase of the column and the more polar mobile phase. A column packed with alkyl-chain modified silica-based sorbent is often used to effect associations with the hydrophobic regions of peptides [115]. The peptides are eluted by applying a gradient of increasing organic solvents in the mobile phase over a period of time. When the percentage of organic solvent is sufficient enough to shift the equilibrium of the peptide towards the mobile phase, the peptide will elute. Subtle differences in conformation and hydrophobicity of the peptides will determine the movement between the two phases and consequently, the rate at which the peptides will move through the column [116-118].

The mobile phase conditions can be manipulated to maximise the separation capabilities of the column. Better resolution of separation can be achieved by varying the composition, pH and ionic strength of the mobile phase. The mobile phase usually contains an ion-pairing agent such as trifluoroacetic acid (TFA) that is used to adjust the pH, solubilize the peptides and minimise any ionic interactions between the peptides and the stationary phase [115]. The addition of TFA assists high-resolution separation and as it is volatile it can be easily removed [117]. The presence of peptides in the eluant may be detected by using an ultraviolet (UV) detector at a wavelength of 214 nm. This wavelength corresponds to the $n \rightarrow \pi^*$ transition of the carbonyl bond [119]. Acetonitrile is commonly used as the organic solvent as it is UV transparent at 214 nm [115].

2.4.2 Peptide Sequence Specific Fragmentation

The primary structure of a peptide can be determined by examining the production of daughter ions formed during CID. Peptides undergo numerous fragmentation processes when analysed by MS/MS. The use of a combination of positive and negative ion MS/MS of peptides provides a more complete set of sequence information.

2.4.2.1 Fragmentation in Positive Ion Mass Spectrometry

Positive ion MS has been the principal means of determining the sequence of peptides. In the positive ion mode, several characteristic fragmentations are observed for linear peptides. The simplest of these fragmentations, the **b** and **y** cleavage patterns are commonly used for sequence determination because they correspond to cleavages of the amide bond (Figure 2.4) [98, 110, 120, 121].

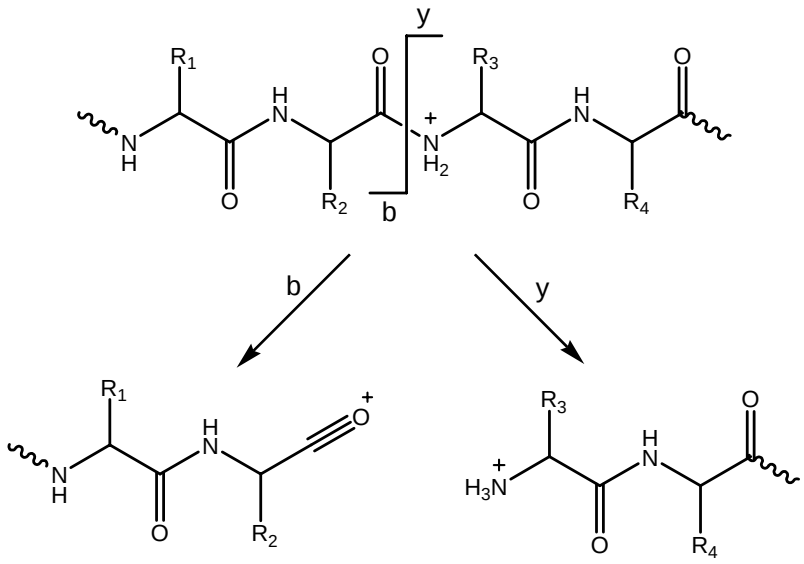


Figure 2.4: Characteristic fragmentation of linear peptides in positive ion MS to form **b** and **y** fragmentation ions [120].

Originally, the **b** fragmentation ions were proposed to be acylium cations (Figure 2.4), however, more recent publications have suggested that these **b** ions exist predominantly as the more stable oxazolones [122-126]. The proposed mechanism for the formation of the cyclic oxazolone **b** ion is illustrated in Figure 2.5. Additional cyclic structures for the **b** fragmentation ion have also been proposed [127, 128].

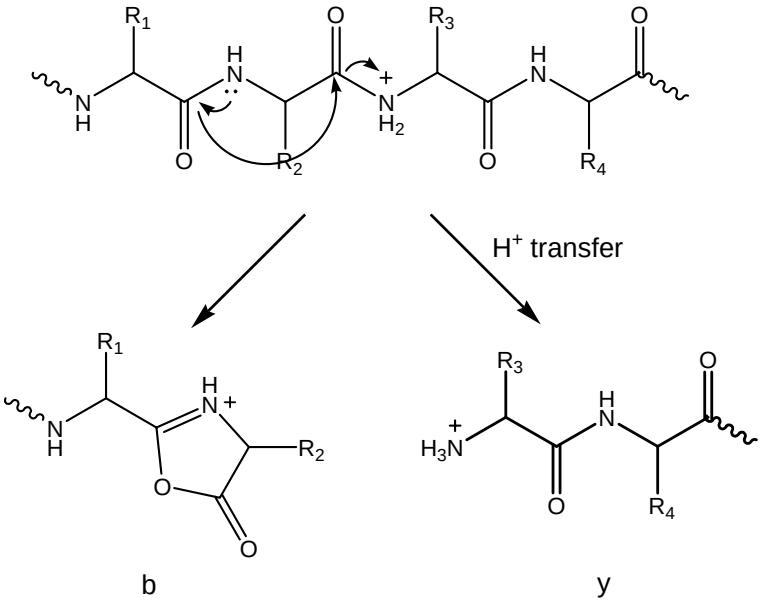


Figure 2.5: Formation of the oxazolone **b** and **y** ions from peptides in positive ion MS [123-126].

Protonated peptides fragment primarily under charge-directed reactions when activated under low energy collision conditions. The proton responsible for initiating fragmentation can be localised at any of the amide bonds and results in an array of peptide bond cleavages that gives a complementary series of **b** and **y** ions. Distinctive mass differences between successive fragment ions of the same type allow for identification of sequential residues (Table A1) [120, 129]. The **b** fragmentation ions provide sequencing information from the C-terminal end of the peptide, whilst the **y** fragmentations allow characterisation from the N-terminal end [130]. One disadvantage of using MS for peptide sequencing is that it is difficult to distinguish between isomeric Leu and Ile (both *m/z* 113) and isobaric Lys and Gln (both *m/z* 128) using low-resolution spectra. In order to differentiate between these residues, Edman sequencing (Section 2.4.4) can be employed, or high-resolution MS and MS/MS for Lys and Gln.

There are some peptides that may not produce desirable spectra in the conditions used for positive ion MS. Examples of these, are peptides containing residues with modified side chains such as sulfated and phosphated Ser, Thr and Tyr residues (Section 2.4.3), as well as acidic residues (eg. Glu and Asp residues). The high acidity of these peptides means that they may form deprotonated ions $(M-H)^-$ more readily than protonated ions $(MH)^+$ [131]. As a consequence, the use of negative ion MS in peptide sequencing may be used to complement information obtained from positive ion examinations.

Fragmentation does not occur in equal abundance along the peptide backbone, as the fragment ions are not equally stable [98]. For example, dominant cleavage often occurs selectively C-terminally to Asp and Glu, as the acidic side chains provide a local mobile proton that is capable of catalysing bond cleavage [98, 132]. A similar induced cleavage has been noted for His [127]. Fragmentation of the peptide bond at the N-terminal side of Pro is common as Pro contains a tertiary nitrogen that is more likely to be protonated [133, 134]. The presence of basic residues in the C-terminus has been shown to favour formation of **y** fragment ions [120]. Additionally, fragmentation between the first two residues at the N-terminal end of the peptide is energetically unfavourable, so the first **b** ion is sometimes not observed [135].

2.4.2.2 Fragmentation in Negative Ion Mass Spectrometry

The use of negative ion MS in peptide sequencing has only been used in limited studies due to the complexity of the resultant spectrum. Many of the resultant cleavages do not correspond to the backbone amide cleavages, which are considered to be most useful for peptide sequencing. Despite this, the sequencing information obtained from negative ion MS is not only complementary to that of positive ion MS [136], but can provide additional information not obtained in the positive ion mode.

The information provided from amide backbone fragmentation is analogous to the **b** and **y** cleavages, although being formed from entirely different mechanisms and have been named α and β cleavages. The α cleavages provide sequence information from the N-terminal (analogous to **y** cleavages), whilst the β cleavages provide information from the C-terminal (analogous to **b** cleavages) [136]. The mechanism of α and β ion formation is illustrated in Figure 2.6 and has been proposed to proceed via enolate anions formed either in the initial deprotonation or by proton transfer from backbone atoms to a CO_2^- or a $CONH^-$ moiety.

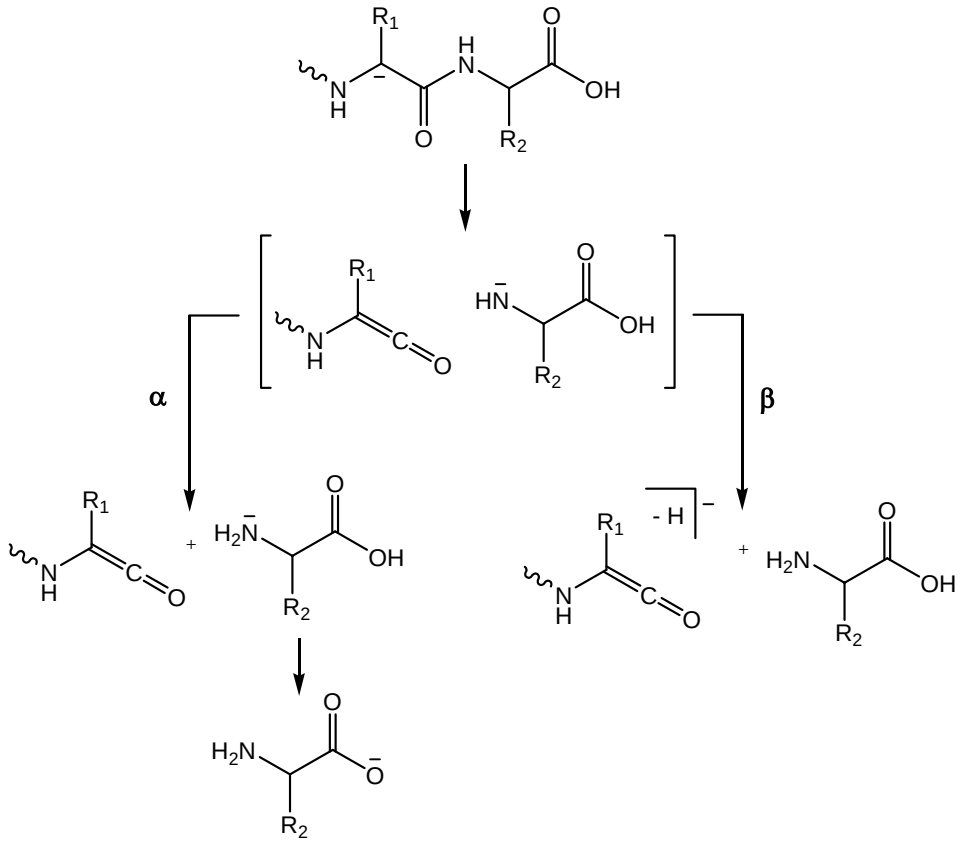


Figure 2.6: Characteristic fragmentation of the peptide amide backbone in negative ion MS to form **α** and **β** fragmentation ions [136].

An alternative structure for the **β** anions has been proposed (Figure 2.7) [137]. This structure is analogous to the proposed cyclic structures of **b** cations formed by decomposition of the peptide $(MH)^+$ ions [126, 128].

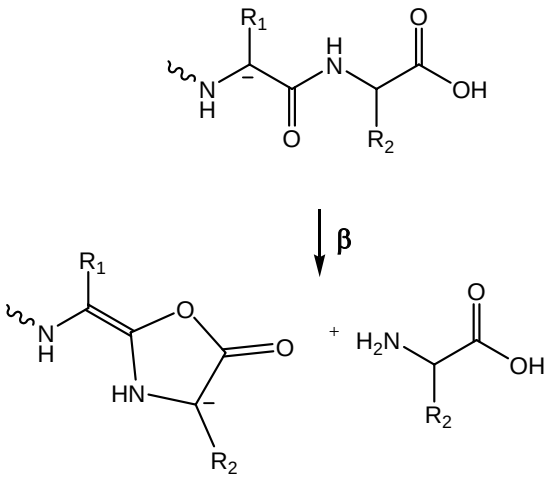


Figure 2.7: An alternative structure for the **β** fragmentation ions formed in negative ion MS [137].

Another backbone cleavage process, the β´ process, can also occur in the peptide irrespective of the C-terminal group (CO_2^- or $CONH^-$) [138]. This process only occurs following cleavage of the amide bond with simultaneous migration of, for example the NH moiety from the C-terminal $CONH^-$ group. β´ fragmentation is initiated by the $CONH^-$ group undergoing nucleophilic addition to the carbonyl group of any amino acid along the flexible backbone, forming a cyclic structure. A subsequent backbone cleavage occurs, which in later steps is analogous to the β cleavage process [139]. This type of backbone cleavage tends to be observed in peptides that lack residues that can undergo facile side chain fragmentations such as Ser, Thr, Asp, Asn, Glu or Gln residues and show strong α cleavages but only minor β cleavages [136].

In addition to the amide backbone cleavages mentioned, a second type of backbone cleavage is observed which is initiated by the side chains of certain amino acids. This occurs for amino acids whose side chains can be deprotonated adjacent to the backbone to form a resonance stabilised anion, such as in Asp, Asn, Phe, Tyr, His and Trp [136]. Generally, as the size of the peptide increases, the occurrence of side chain induced backbone fragmentations of Phe, Tyr and His decreases, while for Asp and Asn the fragmentations are still abundant [123, 136]. These side chain-induced backbone fragmentations have been named δ and γ cleavages. The mechanism of formation of these fragmentation ions are summarised by Bowie *et al*. [136].

In addition to the backbone fragmentations, characteristic side chain fragmentations of many of the common amino acids can occur (Table A2). Many of these fragmentations are observed in the negative ion spectra of very small peptides, however in peptides of four or more amino acids, only several of these side chain fragmentations are prevalent (illustrated in Table A2 in bold). These include; the losses of H_2O from Asp and Glu, the losses of NH_3 from Asn and Gln and the losses of CH_2O and MeCHO from Ser and Thr. The losses from Asp, Ser and Thr side chains often dominate fragmentation of the peptide $[M-H]^-$ anions [136].

Despite its complexity, we use negative ion MS routinely for identifying of peptide primary structure when the positive ion spectrum does not provide complete sequencing information. It has been noted that for small peptides of five to ten residues, positive ion spectra can be ambiguous or incomplete [140], and a combination of both positive and

negative ion spectra are necessary for sequence determination. Negative ion spectra can also be used to distinguish between isobaric Gln and Lys residues that exhibit the same integral mass loss (*m/z* 128) in positive ion mode. In negative ion spectra, the presence of γ and δ cleavage ions from Gln residues readily differentiates this residue from Lys, which does not initiate side chain-induced backbone cleavages [139]. The identification of sulfated and phosphated Ser, Thr and Tyr residues is also possible using negative ion spectra (Section 2.4.3).

2.4.3 Post-Translational Modifications

Post-translational modifications are processing events that result in the change of the structure and properties of the peptide or protein after it has been synthesised. These modifications include the addition of a functional group by acetylation, formation of disulfide bonds, N-terminal pGlu formation and C-terminal amidation [141, 142]. Post-translation modifications are discussed in more detail in Section 7.1.3.2. Many of these modifications result in poor sequence determination by Edman sequencing. The majority of these covalent modifications change the molecular weight of the residue involved, thus MS has provided a useful method for their detection and identification [143].

Sulfation of Tyr residues is an abundant post-translational modification in biological systems and compounds [144]. Under the acidic conditions commonly used in positive ion mode the sulfate group is cleaved from the Tyr residue. As a consequence, the major peak of the spectrum is $(MH)^+ - SO_3$. The presence of the sulfated Tyr residue can be ascertained using negative ion MS, where both the $[M-H]^-$ and $[(M-H)^- - SO_3]^-$ are observed in the spectrum [145]. Similarly, phosphate groups on Ser, Thr and Tyr can also be detected in negative ion MS [146].

2.4.4 Edman Sequencing

One of the limitations of MS in peptide sequencing is its extreme difficulty in distinguishing between isomeric Leu and Ile residues (*m/z* 113). There are ways of doing this but the MS methods are time consuming and not always reliable. As a consequence, Edman sequencing is often used to provide complementary sequencing information to

support that obtained from MS. Edman sequencing was developed by Pehr Edman in the 1960's [147] and is an effective method of N-terminal identification [120]. The process involves a cycle of chemical reactions in which the successive N-terminal amino acid residues are removed and identified by HPLC [120, 147, 148]. In this process, the C-terminal of the peptide is bound to a solid membrane support in order to control the consecutive chemical reactions. In each cycle, the peptide is treated with phenylisothiocyanate, which reacts with the N-terminal residue to yield a phenylthiocarbamoyl derivative, followed by acid hydrolysis of the amide bond to produce a phenylthiohydantoin (Figure 2.8). The phenylthiohydantoin is subjected to HPLC analysis and identified by comparison with known retention times for each of the common amino acids. The new shorter peptide and its newly exposed N-terminus is then subjected to another cycle [147, 148].

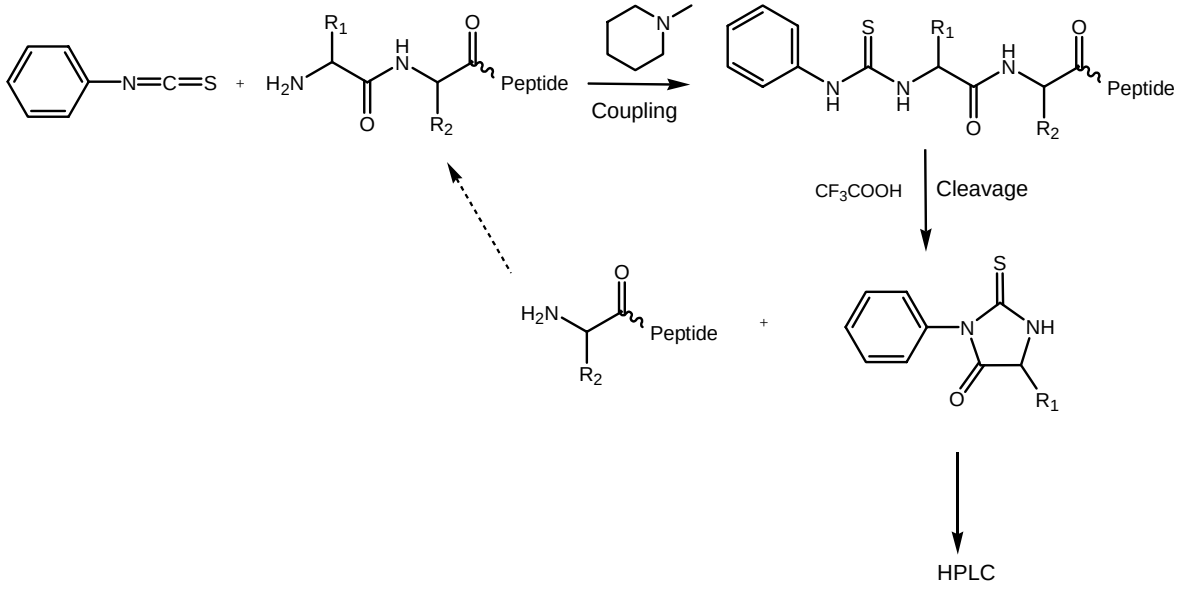


Figure 2.8: Edman sequencing reaction scheme.

Edman sequencing is a very sensitive technique (picomolar range) that is often automated to provide efficient sequencing of peptides of up to 30 amino acids [115]. Alone, care must be taken when using Edman sequencing for primary structure determination [149]. Contamination from incomplete additions of phenylisothiocyanate and degradation reactions can sometimes generate a mixture of products making unambiguous sequencing difficult [120]. Furthermore, gaps are sometimes encountered in positions where residues (Ser, Thr, Arg and His) that generate low yield phenylthiohydrantoin derivatives are present in the sequence [149]. Identification of residues is based on comparison of the known HPLC retention times of common amino acids, thus any uncommon or modified

amino acids will be missed using this approach. Peptides with pGlu residues block Edman sequencing, as the initial phenylthiocarbamoyl derivative cannot form [120]. Similarly, incomplete degradation can sometimes occur when Pro residues are present in a peptide, blocking sequencing [150]. Lastly, there is a risk that as the peptide gets shorter it can be washed away from the solid membrane support causing incomplete sequencing of the C-terminal end. Despite these limitations, Edman sequencing is a valuable technique when coupled with MS in ascertaining the primary structure of peptides.

2.5 ESI-MS of Protein Complexes

The function of many biological molecules depends on its specific, non-covalent interactions with other molecules. Proteins interact with a variety of biological molecules including other proteins, peptides, oligonucleotides (such as DNA and RNA), metal ions and other small molecules. An understanding of these interactions is essential in order to understand the biological functions of these molecules and their role in cellular processes. With the recent demonstration that MS could be used to measure protein complexes as large as 2 MDa, MS has become a popular tool for investigating the non-covalent interactions involved in protein-ligand complexes [111]. ESI-MS has been used extensively as a “soft” ionisation method, which results in the preservation of weak non-covalent interactions allowing intact complexes to be detected and studied.

The folding of proteins, formation of multimers and non-covalent interactions between the proteins and ligands are important for the biological function of these proteins. The primary forces that govern these interactions in solution include hydrogen bonds, hydrophobic effects, salt bridges, van der Waals forces and electrostatic interactions [151]. In order to study the protein complexes formed it is essential that the non-covalent associations are maintained. Under appropriate ESI conditions, it is possible to preserve the intact protein complex upon transition to the gas-phase and retain the important aspects of its solution structure including the non-covalent interactions [151-153]. ESI-MS does not provide the same structural information as nuclear magnetic resonance (NMR) and X-ray crystallography, but ESI-MS can be used to determine important features of proteins and protein complexes including stoichiometry, characterisation of non-covalent interactions and conformational changes upon binding.

Two of the earliest reported studies to show the ability to observe intact non-covalent interactions in ESI-MS were the intact heme-myoglobin complex [154] and a non-covalent complex formed between the immunosuppressive FK binding protein and immunosuppressive agents FK506 and rapamycin [155]. Since these reports, ESI-MS has been used extensively to study weakly bound complexes including protein-protein, protein-metal and protein-oligonucleotide interactions [151]. The large number of examples in the literature has demonstrated that ESI-MS can be utilised for studying complexes in the gas-phase (for reviews see [151, 156]).

The stoichiometry of a complex can be easily obtained from the mass spectrum by the measurement of the molecular weight of the complex [151]. In addition, subunit interactions including the formation of a monomer, dimer or multimer can be determined. The detection of a non-covalent interaction relies solely on the observation of a mass shift in the spectrum from the charge state distribution of the native protein [156]. Changes in charge-state distribution in the ESI spectrum can also be used to probe the behaviour of complexes in solution. Conformational changes in a protein upon binding ligands or metal ions are evident by a change in charge state distribution. The extent of multiple charging of protein ions in ESI-MS often reflects their tertiary structure in solution, as more compact structures cannot include as many charges as partially unfolded isomers [157]. A narrow charge distribution of a low charge state represents the preservation of the higher ordered structure of the native protein complex as fewer charged sites are exposed in a compact structure [151, 157]. As a consequence, a change in charge state distribution may be indicative of a change in conformation.

The strength of interactions, stability, relative binding affinities and structure can be probed by competitive binding studies in which several species are competing for a binding site [158]. These characteristics can also be explored by inducing decomposition of complexes as a result of manipulation of the variables in the atmospheric-vacuum interface region of the mass spectrometer. Non-covalent complexes can be fragile in the gas-phase and may dissociate by the application of excess heat or collision energy in the atmospheric-vacuum interface [158, 159]. Complexes of greater stability generally require greater collision energy and temperature for dissociation, thus providing a measure of the relative structural stability [160] and strength of the interaction. This versatility of ESI-MS

allows the kinetics of biological complexes to be studied making it a desirable tool in the characterisation of biomolecules.

The typical solvent conditions utilised in ESI-MS to achieve maximum sensitivity are unfortunately not optimal for maintaining intact protein complexes. Many protein complexes require physiological conditions to preserve interactions: thus they are highly dependent on conditions such as pH and the ionic strength of the solution. ESI-MS is often intolerant of high salt and buffer conditions and as a result, studies of protein complexes are often done in the absence of salts. Organic solvents may also be added to solutions to improve sensitivity and signal stability [151, 156]. Slow flow rates are often used as they result in the generation of smaller droplets, causing in an increase in signal levels and more sensitive detection [111]. Generally, there is a very fine line between sufficient desolvation of the gas-phase complex and dissociation of that complex, since most gas-phase complexes are fragile. As a consequence, the success of ESI-MS studies of protein complexes depends on carefully controlled conditions that allow the intact complexes to be retained.

There is controversy concerning whether the observations of protein complexes in gas-phase reflect the behaviour of the complex in solution. It has not as yet, been demonstrated conclusively that the activity, binding affinities and conformation of a gas-phase complex is identical to its solution complex [156]. Irrespective of this controversy, it is generally considered that there is a good correlation observed between complexes in the gas-phase and their counterparts in solution [161]. A positive correlation between the three-dimensional (3D) structure of proteins in the gas-phase and solution has been indirectly demonstrated through the ability to differentiate between the denatured protein and the native protein in ESI-MS [162, 163]. This suggests that protein conformation may be conserved during the ESI process. In addition, the stability of the gas-phase complex has been shown to be comparable to the corresponding solution-phase dissociation constants for several proteins [156, 164], and complex studies have shown a correlation between relative binding affinities determined in the gas-phase and in solution [165]. Furthermore, high consistency is also observed in the determination of stoichiometry. For example, CaM binds four Ca^{2+} ions in the gas-phase and in solution [166-169].

Despite the correlations discussed above, several problems have been encountered in ESI-MS investigations. The primary forces that govern non-covalent interactions in aqueous solutions may in some instances play only a limited role in gas-phase measurements [111, 151]. Transition from solution to a gas-phase environment has been shown to strengthen the electrostatic interactions present in some complexes, thus complexes that are primarily held together by electrostatic interactions will be extremely stable in the gas-phase and may not necessarily represent the complex structure in solution [161, 170]. Additionally, in the gas-phase the hydrophobic interactions that are present in the complex in solution have been shown to be weakened due to solvent evaporation [111]. This results in poor correlation between the gas-phase and solution complexes with a high proportion of hydrophobic interactions. For example, competitive binding studies performed by Robinson *et al*. were unable to find a correlation between the dissociation constants of the binding of acyl CoA derivatives to acyl CoA binding protein in solution to the non-covalent complexes in ESI-MS [171].

Differences observed between gas-phase and solution studies in some examples have demonstrated that the information obtained from ESI-MS studies needs to be interpreted with caution. These studies should be used in combination with solution studies to ensure that a correlation between the complex in the two phases is observed.

Chapter 3 Methodology II - Nuclear Magnetic Resonance Spectroscopy

3.1 Secondary Structure Determination of Peptides by NMR

Since the completion of the human genome project, scientists have moved to investigating the proteins encoded by these genes. To completely understand the function and mechanism of action of proteins, it is necessary to identify their 3D structure. The principal technique used to examine the structure of proteins has been X-ray crystallography, which reveals the solid state conformer and circular dichroism that allows the presence of secondary structure to be determined [148]. Recently, the development of NMR spectroscopy has outshone these techniques as the principal structure determination method for proteins in solution [172].

In recent years, NMR has proved, in some cases, to be more advantageous over X-ray crystallography for structure determination. Both techniques allow the 3D structure of a protein to be ascertained at atomic resolution [172, 173]. However, the protein structures determined by X-ray crystallography are dependent on the ability of the protein to form a single crystal and to have sufficient resolution of the crystal diffraction [173]. Furthermore, in the crystal structure, significant functional differences may exist compared to the solution structure due to additional forces that are found in the solid state, such as crystal-packing forces that are not present in the species when in solution. In contrast, the solution structures of proteins determined from NMR spectroscopy are obtained under near physiological conditions and thus may be more relevant to studying protein function and mechanism of action. In addition, the dynamic nature of proteins is also more amenable to be studied by NMR [173, 174].

Although NMR spectroscopy is the favoured method of structure determination in solution, it does have some limitations. To be able to determine the 3D structure, the proteins must be soluble and not aggregate readily in solution at millimolar concentrations. NMR spectroscopy is also limited by the size of the protein, but recent development of multi-dimensional experiments and the use of high-field spectrometers, has allowed larger proteins with molecular masses of greater than 30 kDa to be studied [172, 173, 175].

NMR spectroscopy can provide insight into the mechanisms of action of biologically active proteins and peptides by providing information on dynamics, conformational equilibria, molecular structure and intra- and intermolecular interactions [172]. This chapter describes the basic principles of NMR spectroscopy and its application to the 3D structure determination of peptides in solution.

3.2 Nuclear Magnetic Resonance Spectroscopy

NMR experiments are possible as some nuclei behave as bar magnets due to the presence of an intrinsic angular momentum, or a spin, described by the spin quantum number, I. Only nuclei with non-zero spin quantum numbers, in particular ^{1}H, ^{13}C and ^{15}N (I = ½) nuclei, are NMR active and give rise to NMR signals. When NMR active nuclei are subjected to an external applied magnetic field $\mathbf{B_0}$, the nuclei spins orient themselves in (2I + 1) ways with respect to $\mathbf{B_0}$ [119]. ^{1}H, ^{13}C and ^{15}N spins can orient themselves in two possible orientations; a low energy orientation parallel to the external magnetic field $\mathbf{B_0}$ and a high energy orientation that is antiparallel [172]. Nuclei can move from one energy state to the other by either absorbing or emitting electromagnetic radiation corresponding to the energy difference between the states [119].

The difference in population of the two orientations is responsible for the observed NMR signal and is usually very small, as determined by the Boltzmann distribution (3) [119].

$$\frac{N_\beta}{N_\alpha} = e^{-\Delta E/kT} \qquad \textbf{(3)}$$

where k is the Boltzmann constant, T (K) is the temperature and ΔE is the energy difference between the two spin states N_α and N_β. For example, in a magnetic strength of 14.1 Tesla, the difference in energy for a proton is only in the order of 1 in 10^6, making NMR spectroscopy a relatively insensitive technique [119, 176]. The nuclei can be promoted from the N_α state to the higher energy N_β state when irradiated by an applied *rf* pulse that corresponds to the energy difference between the two states and matches the frequency at which the spins naturally precess in the magnetic field. This corresponds to the Larmor frequency and is defined by:

$$\nu = \frac{\gamma B_0}{2\pi} \quad \textbf{(4)}$$

where γ is the gyromagnetic ratio (a proportionality constant differing for each nucleus) and B_0 is the strength of the applied magnetic field. For a magnetic field of 14.1 Tesla, the Larmor frequency corresponds to 600 MHz for ^{1}H nuclei [119].

As stated above, a slightly larger population exists in the lower energy state, creating a net magnetisation **M,** which is oriented parallel to $\mathbf{B_0}$ along the *z* axis (Figure 3.1). When irradiated by an applied *rf* pulse, a change in population of the orientations will result, creating a change in **M** [176]. Commonly, an *rf* pulse with a length of time sufficient to rotate **M** through 90°, is applied along the *x* axis to the equilibrium state, which flips **M** into the *xy* plane [172]. An oscillating magnetic field $\mathbf{B_1}$ along the *x* axis is produced. The oscillating magnetic field interacts with **M** to move the magnetisation towards the *xy* plane. The receiver coils are positioned along the *x* and *y* axis [119, 176, 177].

When the *rf* pulse of μs length is turned off, **M** relaxes to equilibrium (*z* axis) over time [176]. There are two relaxation processes that can occur; longitudinal relaxation and transverse relaxation. Longitudinal relaxation involves a transfer of energy from the excited protons to those nuclear spins around it that are tumbling at the suitable frequencies. The result is an increase in *z* magnetisation as **M** returns to the *z* axis [177]. Alternatively, transverse relaxation involves energy transfer among the precessing protons and a decrease in *x* and *y* magnetisation is observed. The resultant effect is characterised by dephasing, line broadening and signal loss [177]. The signal obtained from the decay period of **M** back to equilibrium in the *xy* plane is called free induction decay (FID). Fourier transformation of this FID converts it from a function of time to a function of frequency, resulting in a one-dimensional (1D) NMR spectrum [172].

For a given nucleus, the actual magnetic field present at the nucleus is dependent upon the nature of the surrounding electronic environment, which is termed shielding. As a result, chemically distinct nuclei resonate at different frequencies. Variations in the resonance frequency due to shielding are referred to as the chemical shift. Hence, chemical shift values are dependent on the magnetic environment of the particular nucleus [178].

Information about the molecular structure can be obtained from 1D NMR spectra in the form of chemical shifts, coupling constants, multiplicity of signals and integration. As the size and complexity of the molecule increases, however, the overlap of resonances also increases. This means that for larger molecules such as peptides and proteins, spectral information must be obtained from more complex NMR experiments.

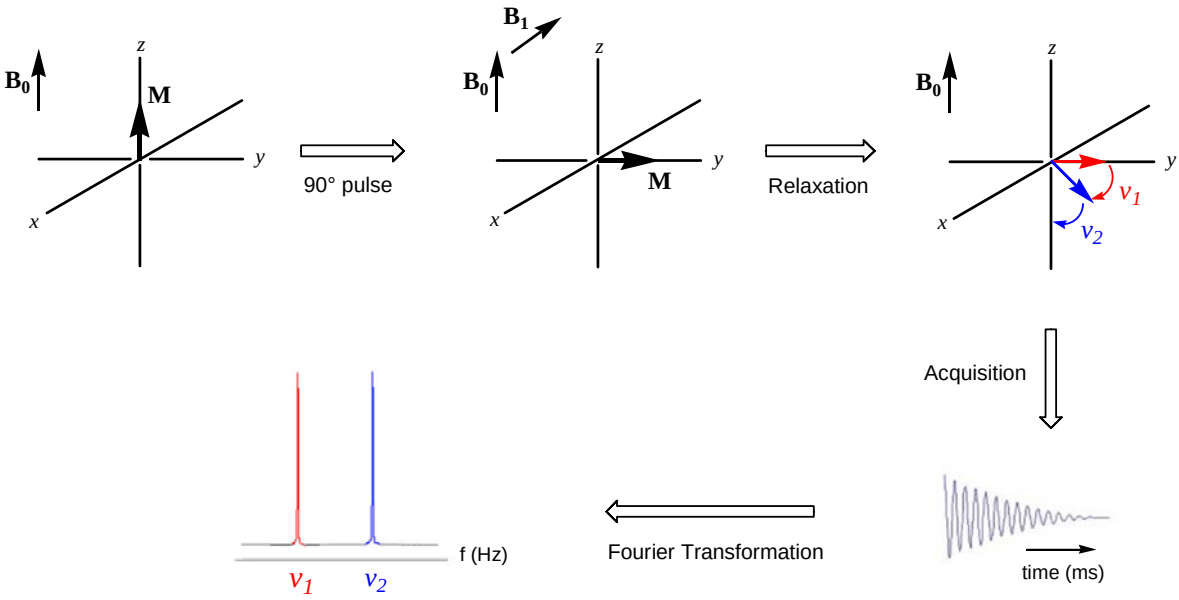


Figure 3.1: Representation of the 1D NMR experiment. The vector diagrams representing the precession of resonances following the application of a 90° *rf* pulse and the 1D spectrum is obtained from the Fourier transformation of the FID in the frequency (f) domain.

3.3 Two Dimensional NMR and Resonance Assignment

During the past ten years NMR has developed into a powerful method for peptide and small protein (100 amino acids or less) 3D structure determination [179]. The development of multi-dimensional experiments has provided a way to resolve the large amount of resonance overlap seen in 1D NMR spectra of peptides. Typically, two-dimensional (2D) NMR experiments consist of: a preparation period, in which the magnetisation is changed from its equilibrium state; an evolution period t_1, where the chemical shifts of individual nuclei will be observed; a mixing period τ_m, during which the spins are correlated to each other; and a detection period t_2, in which the frequency is measured [119, 172]. Multi-dimensional experiments are generally recorded with sequentially incremented lengths of the pulse sequence introduced into the evolution period t_1 to produce a data matrix $s(t_1,t_2)$

[180]. Fourier transformation of this data matrix yields the 2D frequency spectrum $S(\omega_1,\omega_2)$ (Figure 3.2) [174]. The assignment of all the resonances in the peptide is possible through the use of a combination of 2D experiments that are discussed below. This provides the foundation for secondary structure determination.

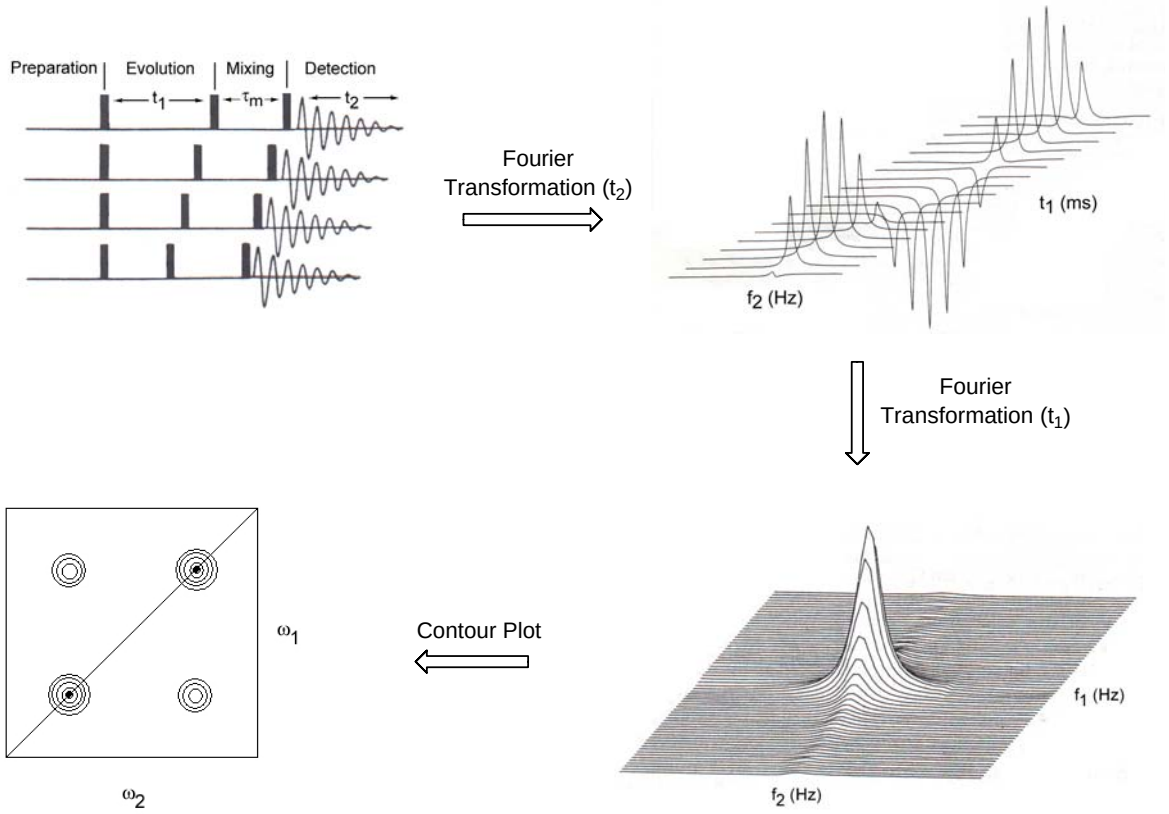


Figure 3.2: Schematic representation of a generalised 2D NMR experiment. Adapted from [179, 181].

3.3.1 Correlated Spectroscopy

Correlated spectroscopy (COSY) is an NMR experiment that provides information about protons that are spin-spin coupled to each other, such as protons on adjacent nuclei. It is a useful experiment for assigning proton resonances within specific spin systems or the same amino acid residue [27]. COSY is the simplest 2D NMR experiment, consisting of two *rf* pulses that are spaced by a variable time t_1 (5) [181]. The pulse sequence is repeated with incremented intervals of t_1. Fourier transformation of data collected after the second pulse time t_2, produces the 2D NMR spectra.

$$90° - t_1 - 90° - t_2 \qquad \textbf{(5)}$$

The magnetisation evolves with its chemical shift in both time periods. The second *rf* pulse rotates the *y* axis magnetisation component through 90° to the –*z* axis and the magnetisation is redistributed by coherence transfer among the coupled nuclei [181]. After Fourier transformations of the t_1 and t_2 dimensions, the resultant spectrum displays cross-peaks symmetrically about the diagonal, indicating mutually spin-spin coupled protons. The couplings are limited to two or three bonds and are often dependent on bond character, thus cross-peaks only occur between protons in the same amino acid residue [175].

Modifications have been made to the COSY experiment to improve phase sensitivity, allowing natural line width and shape of the peaks to be obtained in both dimensions. The consequent improvement of resolution is necessary for the study of larger biological molecules such as peptides and proteins [175]. In conventional COSY experiments, the diagonal cross-peaks are 90° out of phase with each other, resulting in dispersion of the diagonal when the cross-peaks are phased for maximum resolution. By inserting an extra pulse into the pulse sequence and rotating the pulses through an adequate number of phase cycles, the improved double-quantum filtered (DQF) COSY can be obtained [175]. In DQF-COSY spectra, minimal dispersion of the diagonal is seen. However, by improving the diagonal resolution in order to prevent obscuring cross-peaks, a loss in sensitivity by a factor of two relative to traditional COSY experiments, is observed. The lower signal-to-noise ratio has minimal disadvantages as the spectrum is simpler and the cross-peak intensities are increased relative to the diagonal peaks [182].

3.3.2 Total Correlation Spectroscopy

The total correlation spectroscopy (TOCSY) experiment, also referred to as the homonuclear Hartmann-Hahn (HOHAHA) experiment, gives rise to cross-peaks that correlates all resonances in the same spin system [27]. As coupling cannot occur through the amide bond, the same spin system corresponds to the individual amino acid residues for the peptide [119]. Each spin system can be identified as a particular amino acid based on characteristic patterns in the cross section of a TOCSY spectrum [174, 181].

TOCSY experiments utilise cross-polarisation resulting from the addition of a spin-lock field τ_m to the pulse sequence. The pulse sequence of the TOCSY experiment is illustrated in (6) [181].

$$90° - t_1 - \tau_m - t_2 \quad \textbf{(6)}$$

The spin-lock field is a single coherent *rf* field that causes an oscillatory exchange of the magnetisation of two spins when these spins have identical effective *rf* fields [181, 183]. During the application of the spin-lock field, the two spins will temporarily become equivalent. If considering a two spin system, a complete exchange of magnetisation will occur when spin-lock mixing time is equal to ½J, where J is the scalar coupling constant between the two spins [181, 184]. In larger spin systems such as in biological molecules, the spectra are usually obtained from a variety of spin-lock mixing times to ensure that a complete exchange of magnetisation has occurred throughout the entire spin system [181].

TOCSY experiments can offer higher sensitivity and resolution than COSY experiments [182] as all the peak components in the TOCSY spectrum are in phase and absorptive [179]. However, negative peaks can occur in the spectrum due to nuclear Overhauser effects (NOE). This is overcome by applying “clean” TOCSY pulse sequences that inhibit the occurrences of cross-relaxation peaks, but this is generally not a problem when short spin-lock mixing times are used [185].

3.3.3 Heteronuclear Correlation Spectroscopy

Information about coupling between protons and heteronuclei (eg. ^{13}C and ^{15}N) are provided by heteronuclear correlation spectroscopy. The pulse sequence for the simplest heteronuclear correlation experiment involving coherence transfer of magnetisation between a proton (I spin) and a heteronucleus (S spin) is illustrated in (7) [181].

$$\begin{aligned} &\text{I:} \quad 90° - t_1 - 90° \\ &\text{S:} \qquad\quad 90° - t_2 \end{aligned} \quad \textbf{(7)}$$

2D spectra are produced with the I spin frequencies in one dimension and the S spin frequencies in the other dimension. Cross-peaks indicate that the correlated nuclei are scalar coupled, but in contrast, there are no diagonal peaks. The absence of diagonal peaks occurs because the 90° pulse that is applied to the S spin is not applied before the t_1 time period [181]. Coupling of protons in these experiments can produce multiplets, a problem that interferes with interpretation. A delay period before and after the mixing pulse allows refocusing of the magnetisation before detection and elimination of the multiplicity [119].

The delay period is typically set at $\frac{1}{2}J_{IS}$ for a one bond ^{1}H-^{13}C correlation and the heteronuclear coupling constant J_{IS} is typically in the order of 100 to 200 Hz [119].

Heteronuclear correlation pulse sequences are relatively insensitive for biological molecules because of the detection of low abundance ^{13}C (1.1 %) spin magnetisation [174, 181]. To improve sensitivity for these types of experiments, it is necessary to begin with polarisation of the proton and finish with the detection of the proton [181]. A commonly used heteronuclear experiment is heteronuclear single-quantum coherence (HSQC). This experiment uses ^{1}H indirect detection rather than direct acquisition of the ^{13}C signal and shows significantly improved sensitivity [181]. The resultant spectrum contains cross-peaks, corresponding to coupled ^{1}H and ^{13}C atoms.

3.3.4 Nuclear Overhauser Effect Spectroscopy

Once correlation experiments (COSY and TOCSY) have been used to identify the types of amino acids present, it is necessary to identify their positions in the peptide sequence. This information is obtained from nuclear Overhauser effect spectroscopy (NOESY) experiments, which rely on through-space connectivities [182]. A NOE is the change in intensity of an NMR resonance when a resonance nearby is irradiated. The relative intensity of a NOE cross-peak is proportional to the inverse of the sixth power of the distance between the two protons involved (i.e. $I \propto r^{-6}$) [174]. As a result, NOEs will only be observed if the proton pair is spatially close (< 5 Å). From these experiments it is possible to gain information about the distance between pairs of protons in molecules.

The pulse sequence for the NOESY experiment that allows the detection of NOEs is shown in (8), where τ_m is the mixing time [181, 186].

$$90° - t_1 - 90° - \tau_m - 90° - t_2 \qquad \textbf{(8)}$$

The initial 90° pulse in the sequence creates a transverse *xy* magnetization that causes the spins to precess during t_1, while the second 90° pulse rotates the magnetization along the $-z$ axis. In the mixing period τ_m, magnetisation exchanges between two proximal spins because of cross-relaxation (dipolar interactions) due to the NOE effect [181, 187]. All transverse magnetisation is removed when phase cycling is applied to the system, leaving

only those components in the *z* direction. The final 90° pulse converts the remaining magnetisation into observable magnetisation, which is seen as cross-peaks in the spectra [181, 186].

The resultant 2D NMR spectrum contains cross-peaks that indicate which protons have spatial proximity. As the intensity of the cross-peaks is inversely proportional to the sixth power of the distance between the protons, the intensity of the cross-peak will rapidly decrease as the distance between the two protons is increased. As a result, the maximum distance between two protons that can be detected by NOESY experiments is 5 Å [119].

In any practical application of NOESY experiments, a variety of cross-peaks can be generated by mechanisms aside from cross-relaxation, such as chemical exchange, coherence transfer and spin diffusion. As a consequence, direct measurement of NOE signal intensity is not entirely accurate. Cross-peaks arising from chemical exchange are often not a problem if the chemical structure is known and thus these exchanges are expected. Cross-peaks formed as a result of coherence transfer in weakly coupled systems can be suppressed through the introduction of phase cycling into the pulse sequence [188]. In addition, random incremental variations of τ_m created by a pulse field gradient can result in improvement of the spectrum [181]. Furthermore, spin diffusion resulting from cross-relaxation propagating from one nucleus to the next can affect the cross-peak intensity, but this is commonly only a problem for large (> 10 kDa) molecules and can be decreased by appropriate parameter selection. As a consequence, in NOESY spectra, inter-peak distances are generally defined with upper and lower bounds to correct for these uncertainties in intensity measurement (Section 3.5).

3.4 Sequential Assignment

To be able to analyse data from NMR experiments and enable the 2D structure to be calculated, assignment of all resonances must be completed. Due to the high degree of overlap observed in 1D NMR experiments, these spectra are not useful for assigning resonances. Complete assignment of all resonances in ^{1}H NMR spectra have been made possible due to the use of 2D experiments.

Identification of spin systems of individual amino acids is achieved by the use of COSY and TOCSY experiments. TOCSY experiments can identify all the proton resonances in the same spin system, through in-phase direct, single, double and multiple relay connectivities [189], i.e. cross-peaks from NH to the α, β, γ, δ and ε protons (where present) will be observed. Each spin system can be identified as a particular type of amino acid based on characteristic cross-peak patterns [181, 190]. When overlapping occurs in the TOCSY spectrum, complementary information from the COSY spectrum allows unambiguous assignment. This shows cross-peaks for protons that are coupled through two or three bonds and allow all the protons in an amino acid to be identified. Any remaining ambiguous resonances can be assigned with the aid of the NOESY spectrum.

Once the spin systems have been identified as a type of amino acid, it is necessary to ascertain their position in the sequence by the use of the NOESY spectrum. Generally, for all types of peptide secondary structure, sequential protons will be observable in the NOESY spectrum. The most important connectivities for sequential assignment are NH_i to NH_{i+1}, αH_i to NH_{i+1} and βH_i to NH_{i+1} NOE connectivities (Figure 3.3) [179].

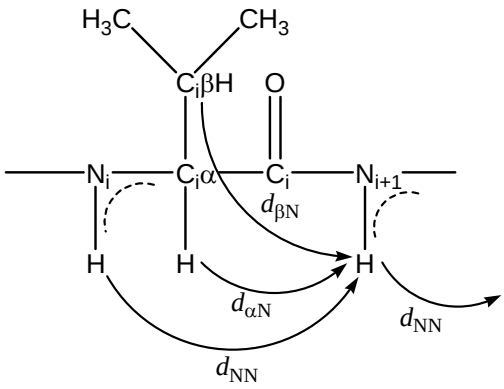


Figure 3.3: Schematic illustration of the NOE connectivities used for sequential assignment of peptide spectra. See Section 3.5 for notation.

Problems can arise for sequential assignment when breaks in sequence are present as a result of coincidental resonances and αH chemical shifts occurring close to the water resonance in aqueous solutions. In addition, the absence of an amide proton in Pro creates breaks in sequential connectivity. However, as the Pro side chain reattaches itself to its backbone nitrogen atom, the δH occupies a similar position to the amide proton in other residues, thus the δH can be used instead for sequential connectivities [175].

Assignment of αC ^{13}C chemical shifts is achieved via ^{13}C-^{1}H HSQC spectra. In the HSQC spectrum, the αH protons are correlated to their attached αC, thus the αC ^{13}C resonances can be identified from the chemical shifts of the cross-peaks in the proton dimension.

3.5 NOE Connectivities

The presence of regions of defined secondary structure can be detected by characteristic patterns of short range (< 5 Å) NOEs in the NOESY spectrum [174, 189]. The standard notation for describing distances between two protons A and B on residues i and j is represented as $d_{AB(i,j)}$. For sequential distances, the residue indices i and j are omitted, for example $d_{AB} \equiv d_{AB(i,j)}$ [174].

The spiral feature of helical structures, inter-residue NOEs can be observed between protons that are three or four residues apart in the sequence. α Helical structures are characterised by strong d_{NN} and medium $d_{\alpha N}$, $d_{\alpha N(i,\ i+3)}$ and $d_{\alpha\beta(i,\ i+3)}$ distances and hence NOEs. In addition, sometimes weak $d_{NN(i,\ i+2)}$ and $d_{\alpha N(i,\ i+4)}$ NOEs can supplement this characteristic NOE pattern. β Strands, on the other hand, are characterised by very strong $d_{\alpha N}$ and weaker d_{NN} NOE connectivities as a result of much more extended structures. The absence of other short range NOEs from NH and αH protons is highly diagnostic for β strand structures [189, 191]. Figure 3.4 illustrates these characteristic NOE connectivities.

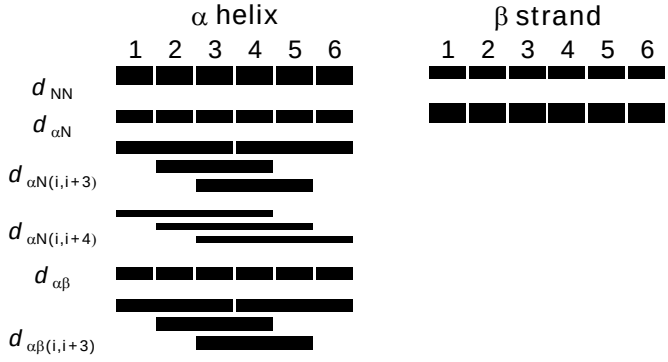


Figure 3.4: Characteristic patterns of short range NOE connectivities seen for ideal α helices and β strands. The relative NOE intensity is indicated by the thickness of the bar [174].

NOE connectivities can also be used to identify *cis-trans* isomerisation of the imide (X-Pro) peptide bond. Variations in the distances between NH_i, αH_i, $\delta CH_{2(i+1)}$ and $\alpha H_{(i+1)}$

are seen in the *cis-trans* isomerisation. The *cis* isomer shows shorter distances in $d_{\alpha\alpha}$ and $d_{\alpha N}$ connectivities relative to the *trans* isomer (Figure 3.5) [174].

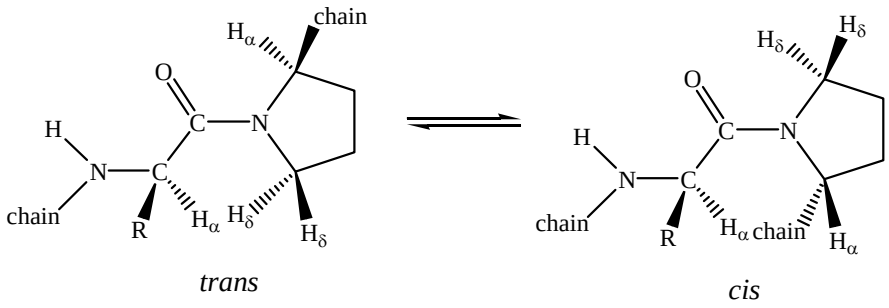


Figure 3.5: The *cis-trans* isomerisation of the imide (X-Pro) peptide bond [174].

The vast majority of peptide bonds in proteins exist as *trans* conformers (relative to αC) as a result of steric hindrance between adjacent αC substituents. However, it is common that the imide (X-Pro) peptide bond is in the *cis* conformer as *cis* and *trans* conformers are relatively close in energy [174, 192]. Both conformers can be observed in NMR simultaneously as the *cis-trans* isomerisation is slow on the NMR time scale [174, 192].

3.6 Secondary Shifts

Chemical shifts are a sensitive measure of molecular composition, conformation and environment. They are influenced by a number of factors including local electronic fields, proximity to aromatic rings, polar and charged groups, local magnetic anisotropies and bond hybridisation [193, 194]. As a consequence, regions of regular secondary structure will have a predictable effect on the chemical shift of nuclei in the peptide. Reference chemical shift values of protons in the twenty common amino acids are obtained empirically from random coil unstructured peptides in water [194]. The variation of the resonances of the protons in the peptide from the corresponding random coil values is referred to as a secondary shift (Δδ). Δδ analysis is only applicable to backbone αH, αC and amide proton resonances. A negative Δδ indicates that the proton resonance is upfield from the random coil chemical shift and a positive Δδ designates a downfield shift. Δδ values for αH and αC are usually smoothed over ± 2 residues to minimise the effects of local chemical shift influences, such as polar and charge moieties and nearby aromatic groups [195]. By plotting the Δδ values against the amino acid sequence, regions of the

peptide that exhibit upfield or downfield shifts can be seen, revealing the presence of defined secondary structure.

Statistical analysis has shown that there are characteristic shifts of backbone atoms associated with certain secondary structures. αH resonances show a large upfield shift of approximately 0.39 ppm when present in a region of α helical secondary structure and when incorporated into β strands exhibit a downfield shift (~ 0.37 ppm). In contrast, αC ^{13}C resonances are shifted downfield in α helices and upfield in β strands [193]. Amide protons are more sensitive to temperature and pH, but a correlation with secondary structure has still been observed. Amide proton resonances are shifted upfield in α helices and downfield in β strands. In addition to the general upfield resonance shift in α helices, the amide protons at the N-terminus are shifted downfield relative to those at the C-terminus [193]. This effect may be a result of helix-dipole deshielding at the positive N-terminal end and shielding at the negative C-terminal end [193, 196]. Furthermore, Δδ values that approach zero can indicate regions of conformational flexibility (i.e. a random coil or structure with little order).

In amphipathic α helices, the amide proton Δδ have been found to vary periodically every three to four residues [197, 198]. These peptides display a difference in the length of the hydrogen bonds along the α helix and as a consequence, the helix is often curved rather than linear [198]. This is characterised by the resonances on the hydrophobic face being shifted downfield and those on the hydrophilic face shifted upfield from the random coil values [197, 198]. The intramolecular hydrogen bonds on the hydrophobic face are strengthened and as a result are shorter, reducing both the solvent access and the potential of intermolecular bonds forming. In contrast, hydrogen bonds on the hydrophilic face are weaker, longer and less linear, resulting in tilting of the carbonyl groups of hydrophilic residues away from the helix in order to form stabilising hydrogen bonds with the solvent [199]. Variation in the amide proton Δδ is due to shorter hydrogen bonds on the hydrophobic face exerting greater deshielding of the amide proton, creating a downfield chemical shift of the amide proton [199].

3.7 Coupling Constants

Information about protons linked by a small number of covalent bonds can be characterised through the spin-spin coupling constants (J values). Coupling constants result from spin pairing of each of the nuclei to the electrons in the bonds between them. As a consequence, each nucleus is sensitive to the orientations of the neighbouring spins. In the NMR spectra, the coupled nuclei will have their signals split into multiplets with the number of signals corresponding to (2I + 1), where I is the nuclear spin of the neighbouring nucleus. The separation of these signals can be measured and is the J coupling constant between the two nuclei [181].

The backbone torsion angles ϕ, defined by $C`_i$, $N_{i+1}$, $\alpha C_{i+1}$ and $C`_{i+1}$ and side chain torsion angles ψ, defined by N_i, αC_i, $C`_i$ and N_{i+1}, are important in determining peptide secondary structure (Figure 3.6) [181]. As the size of the vicinal coupling constant for NH and αH ($^3J_{NH\alpha H}$) is related to the ϕ angle by the Karplus relationship (9), correlation between secondary structure and the $^3J_{NH\alpha H}$ coupling constant (in Hz) can be seen [181, 182, 191].

$$^3J_{NH\alpha H} = 6.4\cos^2(\phi - 60°) - 1.4\cos(\phi - 60°) + 1.9 \quad \textbf{(9)}$$

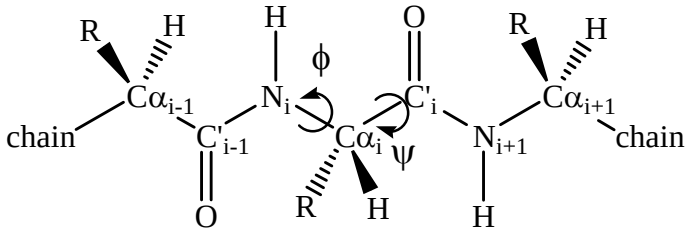


Figure 3.6: The dihedral angles ϕ and ψ. In this representation both ϕ and ψ are equal to 180°.

The Karplus relationship can be represented graphically as in Figure 3.7. Vicinal coupling constants can be obtained with limited accuracy due to localised motion around the peptide N_i - αC_i bond, and as a result, are the average of many measured coupling constants. The ϕ angle can therefore exhibit a range of possible values. It can be seen from Figure 3.7 that for a particular $^3J_{NH\alpha H}$ value, the ϕ angle can have up to four different possible values. To overcome this, ϕ values are restrained over ranges that correspond to particular values. For 5 Hz < $^3J_{NH\alpha H}$ < 6 Hz, ϕ values are restrained to – 60° ± 40°, $^3J_{NH\alpha H}$ < 5 Hz to ϕ = – 60° ± 30°, and $^3J_{NH\alpha H}$ > 8 Hz to ϕ = – 120° ± 40°. When 6 Hz < $^3J_{NH\alpha H}$ < 8 Hz,

ϕ values are not restrained [200, 201]. Typically, in α helices $^{3}J_{NH\alpha H} < 6$ Hz are seen, β sheets are characterised by $^{3}J_{NH\alpha H} > 8$ Hz and $^{3}J_{NH\alpha H} = 6 - 8$ Hz are indicative of random coil structures.

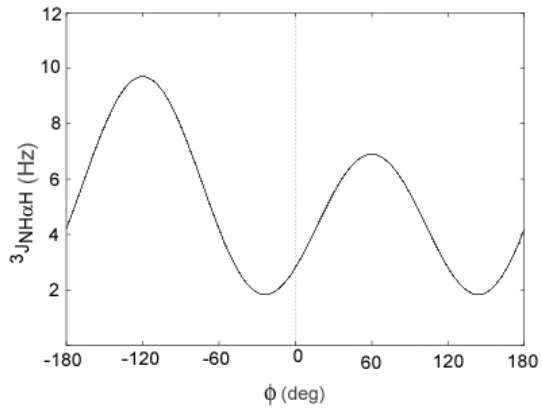


Figure 3.7: Karplus curve illustrating the relationship between the $^{3}J_{NH\alpha H}$ coupling constant and the ϕ dihedral angle. Adapted from [202].

The vicinal coupling constants can be measured directly from the high-resolution 1D ^{1}H NMR spectrum, but as the size of the peptide increases, significant overlap is present, making it difficult to measure these values. An alternative method is to measure the vicinal coupling constants from COSY experiments. The measurement of the antiphase component of the peak separation in the αH and NH cross-peak can lead to the identification of the corresponding coupling constant [203]. However, problems arise as the minimum separation of the COSY antiphase components in the cross-peak is approximately 0.58 times the line width at half of the height. For small coupling constants, as found in α helical structures, the line widths are too broad to determine these values accurately [204].

3.8 Structure Calculations

The 3D structure of a peptide is not directly produced from NMR spectra. Instead, indirect structural information is obtained from which the structure can be determined. Structure determination involves the use of computational calculations of distance restraints derived from the NMR parameters. The structural calculations performed in this thesis were executed using the program ARIA (Ambiguous Restraints for Iterative Assignment, version 1.2) [205, 206] interfaced with CNS Solve (Crystallography and NMR System, version 1.1) [207]. The ARIA program consists of a series of steps that carry out partial

assignment and calibration of the structure restraints that is linked with the organisation of an iterative calculation procedure.

3.8.1 Structure Restraints

Structure constraints are obtained from measurements of the volume or intensity of the NOE cross-peaks in the NOESY spectrum. To a first approximation, the volume of the NOE cross-peaks is inversely proportional to the sixth power of the distance between two protons ($I \propto r^{-6}$) [181]. As a consequence, the intensity of the cross-peaks can be used to estimate the distance r_{ij} between two protons i and j using (10) [208]:

$$r_{ij} = \left[\frac{A(I_{ij})}{I_{ij}} \right]^{-6} \qquad \textbf{(10)}$$

where I_{ij} is the intensity of the cross-peak and $A(I_{ij})$ is the intensity-dependent proportionality factor that is calculated from the average of the ten weakest and ten strongest NOE intensities $A(I_w)$ and $A(I_s)$ respectively.

$$A(I_{ij}) = \left[\frac{I - I_s}{I_w - I_s} \right] \left[A(I_w) - A(I_s) \right] + A(I_s) \qquad \textbf{(11)}$$

where $A(I_w) = (5.0\ \text{Å})^6 I_w$ and $A(I_s) = (1.8\ \text{Å})^6 I_s$

The relationship in (11) assumes that the correlation time is the same for all protons in the molecule and that they are not interacting with any other protons (isolated spin pair). This oversimplifies the situation and may lead to significant deviation from the calibrated values. Calibration of the inter-proton distances by (11) results in an overestimate when the inter-proton distance is less than the reference distances, and an underestimate when the distance is greater than the reference distance [191, 209]. These deviations can be explained by the interference of spin diffusion and local movement within the molecule. The variation from the calibrated values will increase with an increase in correlation and mixing times [181].

The most common reference distance used is that between protons in fixed relative positions, such as the distance between non-degenerate methylene protons, which is fixed

to a distance of 1.7 Å [209]. However, a problem with this can arise, as the chemical shifts of these protons are usually very similar. As a consequence, they correspond to cross-peaks that are close to the diagonal leading to incorrect peak intensities. In addition to this reference, the distance between δ and ε aromatic protons (2.5 Å) [209] is also considered. The aromatic protons however, produce NOEs that are too varied to produce reliable information because of local conformational mobility [181]. The distance between δ and γ protons in Pro can also be used, but the internal motion within the Pro residue can create errors in intensity measurement.

Due to the discrepancies in the reference distances and measurement of the NOE intensities, an alternative approach is used, whereby the distances are estimated directly from the NOE cross-peak intensities using (12).

$$I_{ij} = \alpha r_{ij}^{-6} \qquad \textbf{(12)}$$

where α is the calibration factor. This factor is dependent on properties of the investigated system, in addition to the experimental set up and is determined during the course of calibration [210]. The program ARIA calculates α by comparison of the theoretical (I_{ij}^{th}) and the experimental (I_{ij}^{exp}) (as calculated from the initial ensemble of structures) NOE intensities. This is described by (13).

$$\alpha = \frac{\sum_{ij} I_{ij}^{\exp}}{\sum_{ij} I_{ij}^{th}} \qquad \textbf{(13)}$$

From (13), the target distance for the next iteration can be estimated [210, 211]. In the first iteration where there is no ensemble of structures to calculate I_{ij}^{exp}, α_0 is calculated by assuming that the average distance r_{av} of the spin pairs that cause the NOE is known, as in (14):

$$\alpha_0 = n^{-1} \sum_{ij=1}^{n} \frac{r_{ij}^{\exp}}{r_{av}^{-6}} \qquad \textbf{(14)}$$

It is not possible to obtain exact calibration of restraints due to a number of factors including spin diffusion, internal motion, alternative relaxation pathways and chemical exchange, that can alter the intensity of the signal and in turn the calculated distances [181]. To overcome this, distances are restrained to upper and lower boundaries. The lower boundary is set to a value that corresponds to the sum of the van der Waals radii of two nuclei (1.8 Å) [191] and the upper boundary to the NOE intensity. As a result, NOE cross-peak intensities are generally classified into three categories based on their intensities; weak (1.8 – 5.0 Å), medium (1.8 – 3.3 Å) and strong (1.8 – 2.7 Å) [181, 209].

3.8.2 Ambiguous NOEs

In small molecules, all of the NOE cross-peaks can usually be assigned via chemical shift values. However, as the size of the molecule increases, resonance overlap is observed and the chance of more than one proton having the same chemical shift increases. The resultant effect may result in ambiguity of cross-peak assignment [209]. The quality of structure is dependent on the number of NOE cross-peaks that have been uniquely assigned, thus, the resolution of ambiguous peaks is important for improving the quality of the structure. One method of overcoming this problem is to add a third dimension to the spectra, which will increase the probability of resolving overlapping peaks [212]. However, this method is expensive and time consuming as it involves the introduction of ^{13}C and ^{15}N labelling to the molecule [209].

Another commonly used method is to carry out the preliminary structure calculations using only unambiguously assigned peaks. This enables distances to be set for all proton pairs corresponding to ambiguous NOEs and averaged over the ensemble of structures [209]. Any proton pairs that are distant can be eliminated, but proton pairs that are in close proximity and whose chemical shifts are within a 0.01 ppm range are retained as possibilities [209]. These steps permit the refinement of ambiguous restraints and possibility of new unique assignments.

Alternatively, ambiguous NOE cross-peaks can be handled using the sum averaging method [213, 214]. By the isolated spin pair approximation, the volume of the ambiguous NOE cross-peaks will principally depend on the distance between the two spin pairs as described in (15) [215]:

$$I_{xy} \propto \sum_{a=1}^{N_\delta} r_a^{-6} \qquad \textbf{(15)}$$

where I is the intensity of the cross-peak, a runs through all N_δ contributions to a cross-peak at frequency x and y, and r_a is the distances between the two protons corresponding to the ath contribution. The ambiguous NOE can now be related to an effective distance D by treating the cross-peak as a sum of all of the possible contributing cross-peaks. The resulting distance restraints will account for multiple contributors to the single NOE signal by constraining r_a so that it satisfies (16) [210, 215]:

$$D \propto \left(\sum_{a=1}^{N_\delta} r_a^{-6} \right)^{-1/6} \qquad \textbf{(16)}$$

Although less accurate, this method allows the information contained in an ambiguous NOE signal to be restrained, resulting in improvement of the subsequent calculated structures [213].

3.8.3 Stereo-Specific Assignment

The stereo-specific assignment of resonances in peptides becomes important in defining local conformations. Problems can arise when individual methylene and methyl protons cannot be resolved, for example as in Val and Leu. Assignment of β methylene protons of Leu and γ methyl protons of Val, can be achieved by the use of a combination of $^3J_{\alpha H\beta H}$ coupling constant and a floating chirality method. This method involves assigning protons arbitrarily and using these constraints in the structure calculation procedure. By starting with a small force constant for improper dihedrals and gradually increasing this as the calculations progress, the protons are allowed to ‘float’ between configurations. The resultant structures will exist in the configurations that are consistent with the restraints [179, 216]. Similarly, problems can be associated with assignment of Pro-R and Pro-S resonances in Pro [181, 216].

3.8.4 Restrained Molecular Dynamics and Stimulated Annealing

When calculating peptide structures from NMR-derived restraints, an appropriate method needs to be employed to produce the restrained minimised energy structures. Many methods have been shown to locate local energy minima and result in the generation of inaccurate structures. The most common method is the use of a combination of restrained molecular dynamics (RMD) and dynamical simulated annealing (SA) in Cartesian coordinate space [189].

Molecular dynamics involves the stimulation of the motion of a system of particles, moving according to Newton's laws of motion [191]. RMD attempts to determine the global minimum of a potential energy function that incorporates the experimental NMR restraints. In RMD, the system is allowed to evolve over time from its initial random conformation across a series of positions and velocities. At each step, the force on every atom is determined in accordance to Newton's equation of motion (17) [172, 217].

$$\mathbf{F}_i = \mathrm{m}_i\mathbf{a}_i \qquad \textbf{(17)}$$

where $\mathbf{F}_i$ is the force of a particular atom i, m_i is the mass and $\mathbf{a}_i$ is the acceleration [181]. The derivative of the potential energy V, with respect to the coordinates of the atom can be used to calculate the force $\mathbf{F}_i$ on each atom. With respect to time $\mathrm{t_i}$, the atomic acceleration $\mathbf{a}_i$ can then be expressed as the second derivative of the atomic coordinate $\mathbf{r}_\mathrm{i}$. This derives equation (18):

$$\frac{dV}{d\mathrm{r}_i} = m_i \frac{d^2 r_i}{dt_i^{\,2}} \qquad \textbf{(18)}$$

Equation (18) can be solved to determine the position and velocities of the atoms through successive time steps [181, 217]. The success of the RMD protocol depends on the total potential energy V_{total} function which is used to locate the global minimum and is given by (19) [181, 182, 191]:

$$V_{\mathrm{total}} = V_{\mathrm{bond}} + V_{\mathrm{angle}} + V_{\mathrm{dihedr}} + V_{\mathrm{improper}} + V_{\mathrm{vdW}} + V_{\mathrm{coulomb}} + V_{\mathrm{NMR}} \qquad \textbf{(19)}$$

The potential energies V_{bond}, V_{angle} and V_{dihedr} keep the bond lengths, angles and chirality at their equilibrium values. High energies corresponding to any of these values indicates a

deviation from the ideal values for the bond length, bond angles or dihedral angles. The $V_{improper}$ term is added to retain planarity of the aromatic rings and chirality of the chiral centres. The other physical values are described by V_{vdW} and $V_{coulomb}$; the van der Waals and electrostatic interactions respectively. The final term V_{NMR}, a non-physical term, allows for the inclusion of relevant NMR information and is a way of pulling the protons that exhibit NOE connectivities close to the measured distance r_{ij} that are obtained from NOE intensities. This value also includes information about J couplings through the inclusion of a torsion angle χ_1 term [181, 182]. It is important that these values are chosen carefully to ensure that the experimental restraints used determine the dominant conformation [179, 181, 218, 219]. This is achieved by randomly varying the initial structures. As many structures may be consistent with the experimental restraints, the calculation steps must be repeated multiple times to produce an appropriate ensemble of structures.

The most favourable conformation will have a minimal potential energy function and is determined by employing SA protocol after the RMD calculations. After RMD calculations, SA is preformed to locate the global minimum. This step involves setting the system temperature to a high value (~ 2000 K) to cause an increase in kinetic energy. The kinetic energy is removed by coupling to a theoretical isothermic heat bath [218] and the system is then slowly cooled to overcome local minima energy values and locate the global minimum. The result of this is an improved representation of the peptide structure [172, 179, 219].

Increasing the temperature at which the RMD is performed can enhance the quality of the structural calculations. At higher temperatures, a decrease in experimental constrained violations is observed during the progression of the simulation. This produces calculated structures with lower energy that better resembles the peptide solution structure [191]. In addition, increasing the number of steps in the SA protocol will result in higher convergence of final structures and improved energies [220].

3.9 Structure Quality

Once the calculated structures are generated, it is essential to assess how well the generated structures resemble the solution structure of the peptide. This is achieved by determining

the consistency, precision and accuracy of the calculated ensemble of structures and describes how well the structures satisfy the constraints imposed on them.

How well the calculations satisfy the NMR-derived constraints is expressed in terms of structural violations. Large violations indicate that there is a major problem in the restraints used to generate the calculated structures and these structures do not resemble the actual structure [191]. Violations occur when calculated distances do not agree with the NOE derived restraints. In the final structures, the number of violations should be minimal and below an order of 0.3 Å in size [181]. The precision of the calculated structures is expressed by the root-mean-square derivation (RMSD) of the average structure assembly. When the calculated structures converge and satisfy the NMR restraints, the RMSD will be less than 2 Å and the calculated structures will be close to the solution peptide structure [181]. When errors, indicated by violations and a large RMSD, are encountered after the RMD and SA protocol it can be concluded that there are mistakes in the original resonance assignments of the NMR spectra. This means that the distance restraints need to be reanalysed.

Calculated structures that resemble the actual solution structure will have bond lengths and angles that are closely correlated to the idealised covalent geometry. Furthermore, structures that are consistent will exhibit well-defined backbone dihedral angles [221]. Dihedral angles are assessed by angular order parameters (AOP), which provides a measure of the distribution of the dihedral angles for each residue in the ensemble of calculated structures [221]. AOP is calculated from (20):

$$S(\alpha_i) = \frac{1}{N}\left\| \left(\sum_{j=1}^{N} \alpha_i^j \right) \right\| \qquad \textbf{(20)}$$

where $S(\alpha_i)$ is the order parameter of the dihedral angle α_i of residue i, and j refers to the individual structures from 1 to N [222]. The dihedral angle α_i is expressed as a vector for AOP calculations. For a fully defined angle with no deviation over the ensemble of structures, S will be equal to one. In contrast, where the dihedral angle is randomly distributed, S will be equal to zero [221]. Well-defined residues are generally considered to

have S values greater than 0.9, which corresponds to a standard deviation of the dihedral angle of ± 24° [222].

AOP can be calculated for ϕ and ψ dihedral angles (Figure 3.6) and assesses how well these angles correlate to their ideal values. It has been shown that due to steric interactions, only certain combinations of ϕ and ψ angles are allowed [223]. When the ϕ and ψ angles of peptides with known structures are plotted against each other, they tend to fall into defined regions [224]. Due to their atypical dihedral values, Pro and Gly residues are excluded from these plots. Generally, the regions of the plot with the highest density of residues are considered to be the favourable regions, with over 80 % of residues falling into this region. The residue densities of the other regions progressively decreases and are known as allowed, generous and disallowed regions. This plot is known as the Ramachandran plot (Figure 3.8). The ϕ and ψ angles in residues within different secondary structures will appear in different regions of the plot, hence Ramachandran plots can provide an indication of secondary structure. High quality structures will show well-defined residues in the favourable and allowed regions of a Ramachandran plot [224].

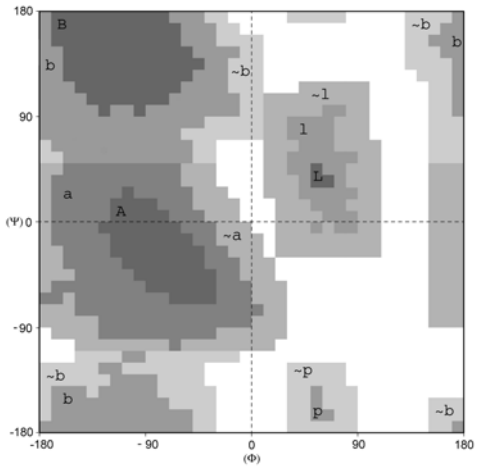


Figure 3.8: The Ramachandran plot illustrating the residue densities of favourable, allowed, generous and disallowed regions. Favourable regions are labelled A and B for α helical or β strand structures respectively, while allowable and generous regions are labelled a and b or ~a and ~b.

3.10 Solvent Selection

In peptide structure determination by NMR spectroscopy, it is important to select an appropriate solvent system that will mimic the different conditions within the body to allow for an accurate physiologically relevant structural representation of the peptide.

Physiological conditions, including pH and ionic strength are obtained through the use of buffers, co-solvent systems and additives such as lipid molecules. As much of peptide and protein action is believed to be at the cellular membrane, it is essential that this is the structural conformer that is studied. Membrane mimicking solvent systems such as various mixtures of water and organic solvents and lipid micelles are commonly used. These solvents induce secondary structure similar to those found in membrane environments [27].

3.10.1 Trifluoroethanol

In water alone, the carbonyl and amide groups of peptides tend to form intermolecular hydrogen bonds with the water molecules. These interactions disrupt the intramolecular hydrogen bonds within the peptide that are responsible for secondary structure [225]. The addition of alcohols such as 2,2,2-trifluoroethanol (TFE) to the solvent favour the formation of intramolecular hydrogen bonds and stabilise the secondary structure [225]. This is due to the larger size, lower polarity and poor basicity of these alcohols [226, 227]. The alcohol is able to displace water from the hydration shell of the peptide, resulting in a favourable increase in entropy [227]. In addition, bulky groups of the alcohols such as the CF_3 group of TFE, sterically hinder the interactions of water with the peptide backbone. Further stabilisation of the peptide structure results from the alcohol's lower relative dielectric constant [228], which leads to a strengthening of charge interactions in the peptide, such as hydrogen bonds and salt bridges. This provides an environment that more closely resembles that of the interior of a protein [225].

Fluoroalcohols such as TFE and hexafluoroisopropanol are more commonly used in co-solvent systems than other conventional alcohols as they display a greater stabilising effect. This effect may be a result of the fluoroalcohol's tendency to form micelle like clusters in aqueous environments [227], producing areas of localised high fluoroalcohol concentration and localised hydrophobic regions, in which intramolecular hydrogen bonds can be strengthened [227]. For larger proteins, these localised hydrophobic regions have been shown to contribute to a weakening of the hydrophobic interactions, which are important in maintaining tertiary structure. Destabilisation of the protein tertiary structure results in unfolding of the protein and as a consequence, little correlation is seen between the structures obtained in aqueous TFE and the native protein form [229, 230]. The amphibian peptides investigated in this study are small and tend to lack tertiary structure,

thus it is reasonable to consider that their structures determined in such co-solvents will resemble that of their native membrane conformers.

There has been concern that the addition of TFE may induce helicity in peptides that are otherwise unstructured in hydrophilic solutions [231]. Sönnichsen *et al*. illustrated that TFE enhanced helicity in peptides only when the peptide displayed a tendency to exhibit helical structure [231]. Stabilisation and induction of β sheets and β turns have also been observed in aqueous TFE solutions [225, 227, 229], demonstrating that TFE does not alone impose helicity in a peptide, but merely enhances this structure if that tendency already exists within the sequence of that peptide.

3.10.2 Micelles

Micelles are spherical aggregates of amphipathic lipid molecules, which form in aqueous solutions when there is a sufficiently high enough concentration; namely, when the concentration is greater than the critical micelle concentration (CMC). The CMC is the minimum concentration of lipid molecules needed to form micelle aggregates in solution [232]. In the micelles, the long, hydrophobic tails of the lipid molecules are directed towards the centre of the aggregate structure forming a hydrophobic core, while the hydrophilic head groups are in contact with the aqueous environment.

Dodecylphosphocholine (DPC) and sodium dodecylsulfate (SDS) are among the most commonly used lipids for studying membrane active peptides and proteins by multi-dimensional NMR experiments. At near neutral pH, DPC has a zwitterionic phosphocholine head group, whilst SDS contains a negatively charged sulfate head group. These molecules form stable micelle structures with small aggregation numbers and low CMC values (DPC: 1 mM and SDS: 8 mM). In addition, these lipids reorient rapidly in solution, thus making them ideal for studying peptides in membrane-like environments in NMR experiments [232].

DPC micelles have a narrow distribution of size, displaying diameters of between 40 and 55 Å [233, 234] and typically containing 55 lipid molecules per micelle [234, 235]. Following the addition of a peptide to solutions of DPC micelles and its subsequent complexation to the micelles, the size of the micelle remains approximately the same,

suggesting that the peptide simply displaces lipid molecules from the micelle aggregate upon binding. A ratio of one peptide to 40 DPC molecules was found to be optimal in the cases of both melittin [234, 236] and glucagon [237, 238].

Peptides can reorient within the micelle and result in an equilibrium between the bound and unbound conformers [239, 240] and two sets of signals are observed in the NMR spectra of such systems. In such a system, a lipid concentration well above that of the CMC is required to ensure that the majority of the peptide is bound to the micelle [232, 234].

Micelles disrupt the protein tertiary structure in a similar manner as TFE [241]. For example, enzymes are rarely active in micelle solutions. The disruption in protein tertiary structure in this system may be due to the severe surface curvature observed in micelle structures [233]. Despite this observation, DPC micelles have been used to successfully determine the solution structures of numerous peptides including melittin [234, 236], glucagon [237, 238] and magainin [242].

Chapter 4 Peptides from the Skin Secretions of *Litoria fallax*

4.1 Introduction

Over 200 species of anurans have been identified in Australia [243]. One of these species, the Eastern Dwarf Tree Frog *Litoria fallax*, is among the smallest of the Australian green tree frogs [243]. This small, slender animal has varied colourings, from uniformed green or brown to green with a pale brown or golden back (Figure 4.1). *L. fallax* has a white, glandular ventral surface with orange thighs, and a distinctive white stripe extending from below the eye to the shoulders. The toes have well developed webbing whilst the fingers display only traces. Females generally range from 25 – 32 mm in size, with the male being slightly smaller at 22 – 26 mm [243].

Figure 4.1: *Litoria fallax*, Eastern Dwarf Tree Frog[*].

L. fallax is a coastal species found in an area ranging from northern Queensland to central New South Wales (Figure 4.2). The animals usually reside in large numbers nearby to water and often among emergent vegetation. They call during the months of October to

[*] image from http://www.banjosretreat.com.au/images/Litoria_fallax

April in a high-pitched, upwards inflecting “wreek-pip-pip”. The first component is very drawn out whilst the latter can be described as short and staccato [243]. *L. fallax* lay small clumps of eggs among submerged vegetation resulting in tadpoles that can easily be mistaken for those from *L. bicolor* [244].

L. fallax is a member of the Dwarf Tree Frog complex, which includes three other species, the Northern Dwarf Tree Frog (*L. bicolor*), the Cooloolah Tree Frog (*L. cooloolensis*), and the Olongburra Frog (*L. olongburensis*) [243, 245]. These species have been found to inhabit some overlapping areas (Figure 4.2).

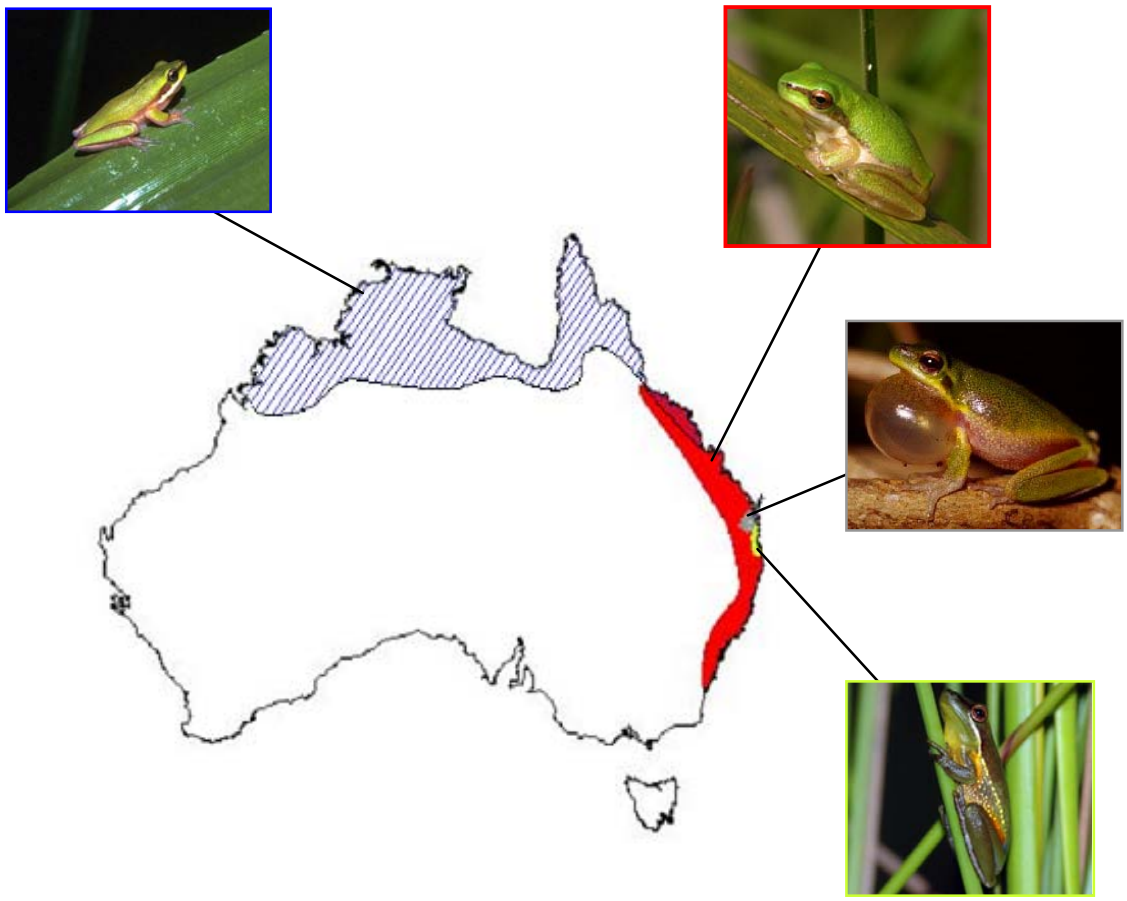


Figure 4.2: The distribution of *Litoria fallax* (red) and related species *L. bicolor* (blue), *L. cooloolensis* (grey) and *L. olongburensis* (green) throughout Australia. Adapted from [243, 245].

L. fallax is the first member of the Dwarf Tree Frog complex that we have studied. The results presented here will detail the isolation, characterisation and biological activity of peptides obtained from the skin secretion of *L. fallax*.

4.2 Results

4.2.1 Isolation of Peptides

Skin secretions of *L. fallax* were collected from seven animals using the surface electrical stimulation method [83] and no variation in the secretion profile of the seven animals was observed. Purification of the secretions was performed by HPLC. A typical HLPC chromatogram is shown in Figure 4.3. Ten peptides were identified from nine separate fractions (labelled A to J in Figure 4.3). There was no peptide material obtained prior to 12 minutes.

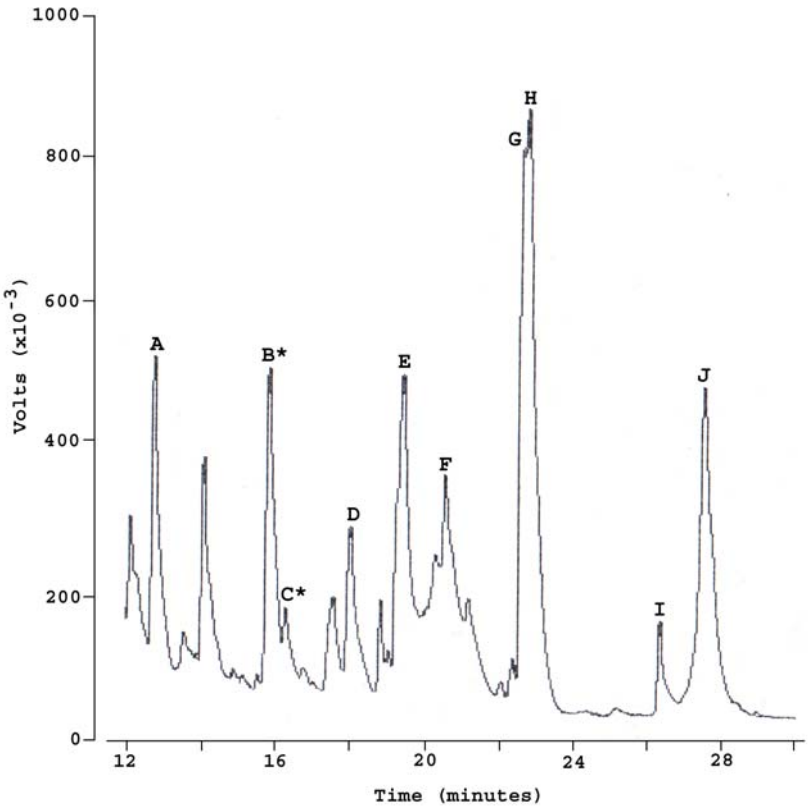


Figure 4.3: Partial chromatogram (12 – 30 minutes) of the HPLC separation of the skin secretion from *Litoria fallax*. As no peptide material was isolated prior to 12 minutes, this part of the chromatogram is not shown. * indicates that the peptide is an enzymatic degradation product. No peptide material was observed in the unlabelled peaks.

The sequences of the peptides secreted by *L. fallax* are listed in Table 4.1 (see Section 4.2.2 for details). The peptides were to be named fallaxins, however due to a peptide isolated from *Leptodactylus fallax* [246] already named this, these peptides have been named fallaxidins and grouped with respect to sequence similarities. The peptide

components isolated in fractions B and C (Figure 4.3), fallaxidin 2.1.1 and 2.2.1, were found to be enzymatic degradation products of fallaxidin 2.1 and 2.2 respectively. Fallaxidin 4 peptides were identified by cDNA cloning (Section 4.2.5) and molecular masses determined from their primary structures were used to locate these peptides in the HPLC fractions. MS sequencing of the fallaxidin 4 peptides will not be discussed here.

Table 4.1: Peptides identified from the skin secretions of *Litoria fallax*. The table indicates the number assigned to the fallaxidin peptides, the sequence, nominal molecular weight (MW) and the HPLC fraction, as indicated in Figure 4.3, from which the peptide was isolated from. Where no HPLC fraction is indicated, the peptide was identified from cDNA precursors (Chapter 7).

Fallaxidin	Sequence	MW	HPLC Fraction
1.1	YFPIPI-NH_2	747	D
1.2	YFPIPF-NH_2	781	E
1.3	YHPF-NH_2	561	A
1.4	YPPI-NH_2	487	
2.1	GLLDLAKHVIGIASKL-NH_2	1645	H
2.1.1	LDLAKHVIGIASKL-NH_2	1475	B
2.2	GLLDFAKHVIGIASKL-NH_2	1679	G
2.2.1	LDFAKHVIGIASKL-NH_2	1509	C
2.3	GLVDFAKHVIGIASKL-NH_2	1505	
3.1	GLLSFLPKVIGVIGHLIHPPS-OH	2191	J
3.2	GLFSFLPKVIGVIGPLIHPPS-OH	2169	
4.1	FWPFM-NH_2	725	F
4.2	FLPLLASLVGGLL-NH_2	1310	I
4.3	FFRVLAKLGKLAK-OH	1518	

4.2.2 Peptide Sequence Determination by Mass Spectrometry

A combination of positive and negative ion ESI-MS was used to provide initial sequencing data. Automated Edman sequencing was then used to distinguish between Leu and Ile residues and to confirm the sequences of all the peptides. The correct residue (Leu or Ile) is shown in all Figures and Tables.

The first of the fallaxidin peptides to be considered is fallaxidin 1.1. The positive ion mass spectrum of fallaxidin 1.1 is shown in Figure 4.4. The **y** fragmentations are indicated below the spectrum and the **b** fragmentations above the spectrum. The sequence of the peptide can be obtained from a combination of the **b** and **y** cleavage information. Fragmentation ions are seen for each of the six residues of the peptide. The loss of 17 Da (NH_3) from the $(MH)^+$ parent ion indicates that the peptide contains a post-translationally modified C-terminal amide. The sequencing data obtained are summarised in Table 4.2.

The negative ion mass spectrum for fallaxidin 1.1 is illustrated in Figure 4.5. The α fragmentation scheme is shown above the spectrum, whilst the β fragmentation ions are listed below the spectrum. Fragmentation ions are apparent for all amide backbone bonds present in the peptide. In addition to the α and β fragmentations observed, side chain-induced backbone cleavages γPhe2 and δPhe2 are also seen. These additional cleavages confirm the presence of a Phe residue at position 2. In the case of fallaxidin 1.1, positive and negative ion spectra provided full sequencing information. Automated Edman sequencing did not distinguish between Ile and Leu because Pro3 and Pro5 hindered the cyclisation/cleavage process [150]. The presence of two Ile residues was determined by cDNA cloning techniques (Section 7.2.3).

The sequencing data for fallaxidin 1.2 and 1.3 are summarised in Table 4.2. Fallaxidin 1.2 is a six residue peptide whose sequence differs from fallaxidin 1.1 by the replacement of Ile for Phe at position 6. Similar sequencing information was obtained from positive and negative ion mass spectra to that obtained for fallaxidin 1.1. In addition to the α and β fragment ions in the negative ion mass spectrum, backbone cleavage ions γPhe2, δPhe2, γPhe6 and δPhe6 were also observed, identifying Phe2 and Phe6. Fallaxidin 1.3 contains four residues and was sequenced from positive ion mass spectra alone. Similarly, its sequence was confirmed by cDNA cloning (Section 7.2.3).

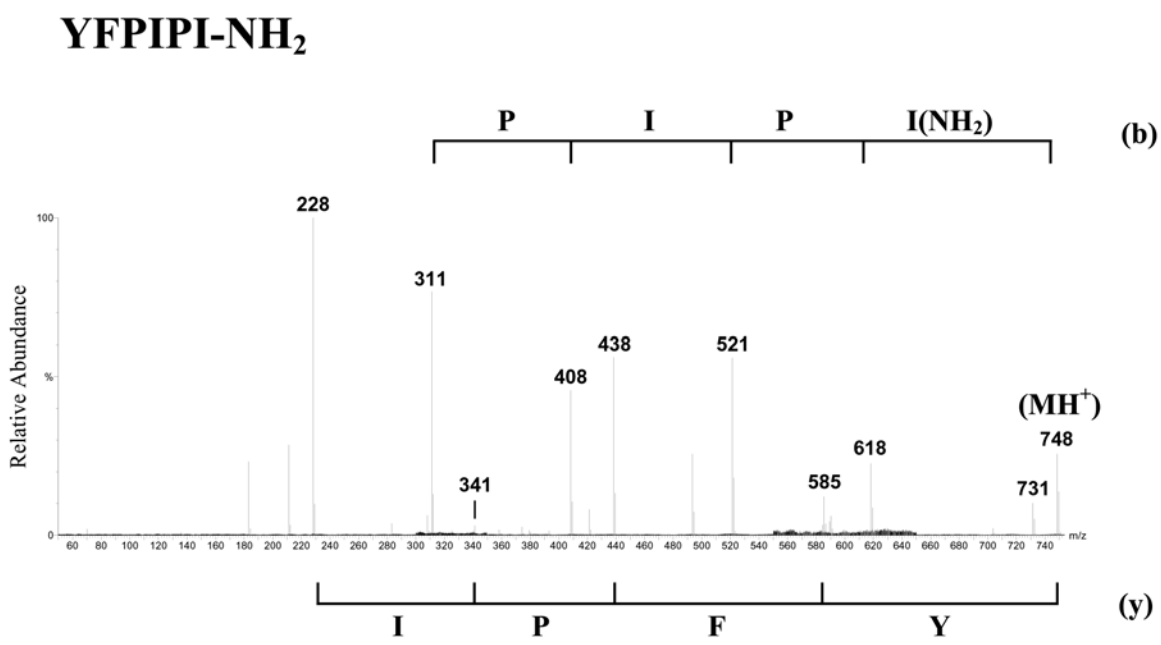


Figure 4.4: Collision-induced MS/MS data for fallaxidin 1.1 illustrating positive ion fragmentations of the $(MH)^+$ species. **b** and **y** cleavage ions are shown schematically above and below the spectrum respectively. Magnification ranges: *m/z* 300 – 350 (x4), 550 – 650 (x6).

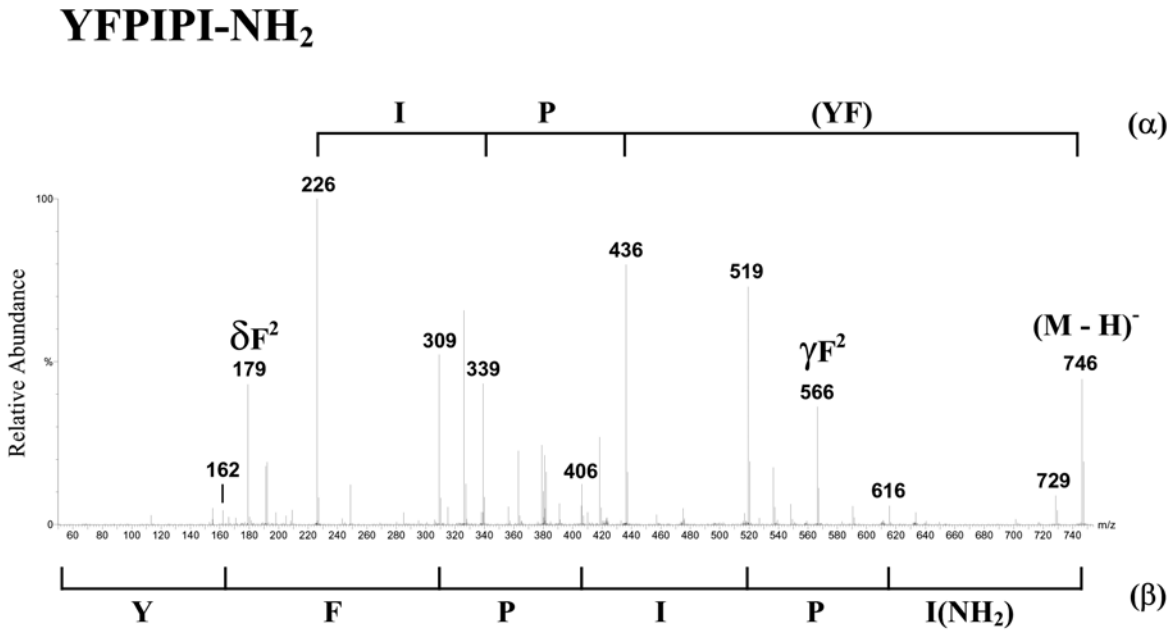


Figure 4.5: Collision-induced MS/MS data for fallaxidin 1.1 illustrating negative ion fragmentations of the $(M-H)^-$ species. α and β cleavage ions are shown schematically above and below the spectrum respectively. Magnification ranges: *m/z* 250 – 400 (x4), 450 – 700 (x3).

Table 4.2: Mass spectral data for fallaxidins 1.1 - 1.3.

Fallaxidin 1.1 YFPIPI-NH_2

$(MH)^+$ *m/z* 748

b ions *m/z* 618, 521, 408, 311 [PIPI-NH_2]

y ions *m/z* 585, 438, 341, 228 [YFPI]

Sequence: YFPIPI-NH_2

$(M-H)^-$ *m/z* 746

α ions *m/z* 436, 339, 226 [(YF)PI]

β ions *m/z* 616, 519, 406, 309, 162 [YFPIPI-NH_2]

γF^2 *m/z* 566, δF^2 *m/z* 179

Sequence: YFPIPI-NH_2

Fallaxidin 1.2 YFPIPF-NH_2

$(MH)^+$ *m/z* 782

b ions *m/z* 618, 521, 408, 311 [PIPF-NH_2]

y ions *m/z* 619, 472, 375, 262, 165 [YFPIP]

Sequence: YFPIPF-NH_2

$(M-H)^-$ *m/z* 780

α ions *m/z* 470, 373, 260 [(YF)PI]

β ions *m/z* 519, 406, 309, 162 [FPI(PF)-NH_2]

γF^2 *m/z* 600, δF^2 *m/z* 179, γF^6 *m/z* 146, δF^6 *m/z* 633

Sequence: YFPIPF-NH_2

Fallaxidin 1.3 YHPF-NH_2

$(MH)^+$ *m/z* 562

b ions *m/z* 545, 301, 164 [YH(PF)-NH_2]

y ions *m/z* 399, 262, 165 [YHP]

Sequence: YHPF-NH_2

The next peptides to be considered are the fallaxidin 2 peptides. Fallaxidin 2.1 varies from fallaxidin 2.2 by only one residue, a Leu5 substitution of a Phe5 residue. The **b** and **y** cleavages in the mass spectrum (Table 4.3) provide the sequence of the peptide except for the orientation of the first two residues at the N-terminus and differentiation between isomeric Leu and Ile and isobaric Lys and Gln. The negative ion mass spectrum (Table 4.3) provides very little additional information on the sequence of fallaxidin 2.1. No Gln δ ions were observed in the spectra, confirming the presence of two Lys residues. Side chain-induced amide backbone cleavages produced γAsp4 ions. Automated Edman sequencing confirmed the primary structure of fallaxidin 2.1.

The positive ion mass spectrum of fallaxidin 2.2 is shown in Figure 4.6, with the **b** and **y** fragmentations illustrated above and below the spectrum respectively. The **b** and **y** fragmentation ions provide amide backbone cleavages for the majority of the residues, however the orientation of the first two N-terminal residues Gly and Leu cannot be determined from this spectrum.

The negative ion mass spectrum of fallaxidin 2.2 (Table 4.3), is dominated by the facile γ cleavage process initiated from Asp4. No α and β cleavages occur from the $(M-H)^-$ parent ion (*m/z* 1678) of fallaxidin 2.2. The absence of Glu γ ions in this spectrum confirms the presence of two Lys residues. Edman sequencing identified the first 15 residues of fallaxidin 2.2 but missed residue 16. Two peptides were synthesised commercially; one containing Leu16(NH_2), the other containing Ile16(NH_2). Comparison of the HPLC data of these two peptides with the native peptide isolated from *L. fallax* identified fallaxidin 2.2 as containing Leu16. This was also confirmed by cDNA sequencing (Section 7.2.3).

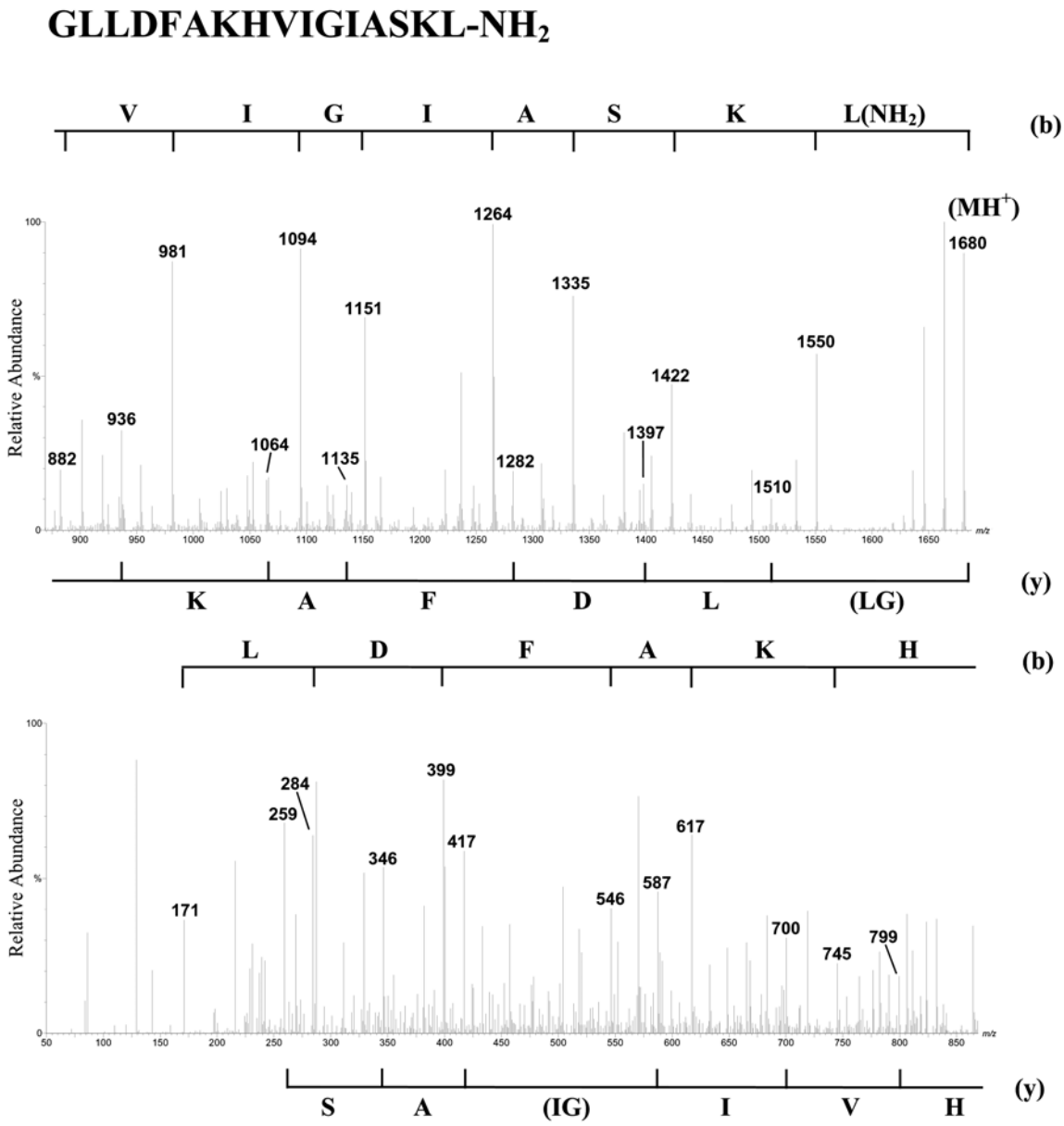


Figure 4.6: Collision-induced MS/MS data for fallaxidin 2.2 illustrating positive ion fragmentations of the $(MH)^+$ species. **b** and **y** cleavage ions are illustrated above and below the spectrum respectively. Magnification ranges: *m/z* 230 – 830 (x4), 900 – 1530 (x2).

Table 4.3: Mass spectral data for fallaxidins 2.1 and 2.2.

Fallaxidin 2.1 GLLDLAKHVIGIASKL-NH_2

$(MH)^+$ *m/z* 1646

b ions *m/z* 1516, 1388, 1301, 1230, 1117, 1060, 947, 848, 711, 583, 512, 399, 284, 171 [LDLAKHVIGIASKL-NH_2]

y ions *m/z* 1476, 1363, 1248, 1135, 1064, 936, 799, 700, 587, 417, 346, 259 [(GL)LDLAKHVI(GI)AS]

Sequence: GLLDLAKHVIGIASKL-NH_2

$(M-H)^-$ *m/z* 1644

γD^4 *m/z* 1344, δD^4 *m/z* 299

[γD^4 - H_2O] *m/z* 1322, [γD^4 - H_2O - CH_2O] *m/z* 1292

Fallaxidin 2.2 GLLDFAKHVIGIASKL-NH_2

$(MH)^+$ *m/z* 1680

b ions *m/z* 1550, 1422, 1335, 1264, 1151, 1094, 981, 882, 745, 617, 546, 399, 284, 171 [LDFAKHVIGIASKL-NH_2]

y ions *m/z* 1510, 1397, 1282, 1135, 1064, 936, 799, 700, 587, 417, 346, 259 [(GL)LDFAKHVI(GI)AS]

Sequence: GLLDFAKHVIGIASKL-NH_2

$(M-H)^-$ *m/z* 1678

γD^4 *m/z* 1378, δD^4 *m/z* 299, γF^5 *m/z* 1263, [γF^5- CH_2O] *m/z* 1233

[γD^4 - H_2O] *m/z* 1360, [γD^4 - H_2O - CH_2O] *m/z* 1330

The final fallaxidin peptide to be considered is fallaxidin 3.1. The positive ion mass spectrum is shown in Figure 4.7. The **b** and **y** cleavage ions are illustrated above and below the spectrum respectively. The **b** and **y** fragmentations provide the full peptide sequence except for the orientation of the last two residues Pro and Ser at the C-terminal end and differentiation between isobaric Lys and Gln and isomeric Leu and Ile.

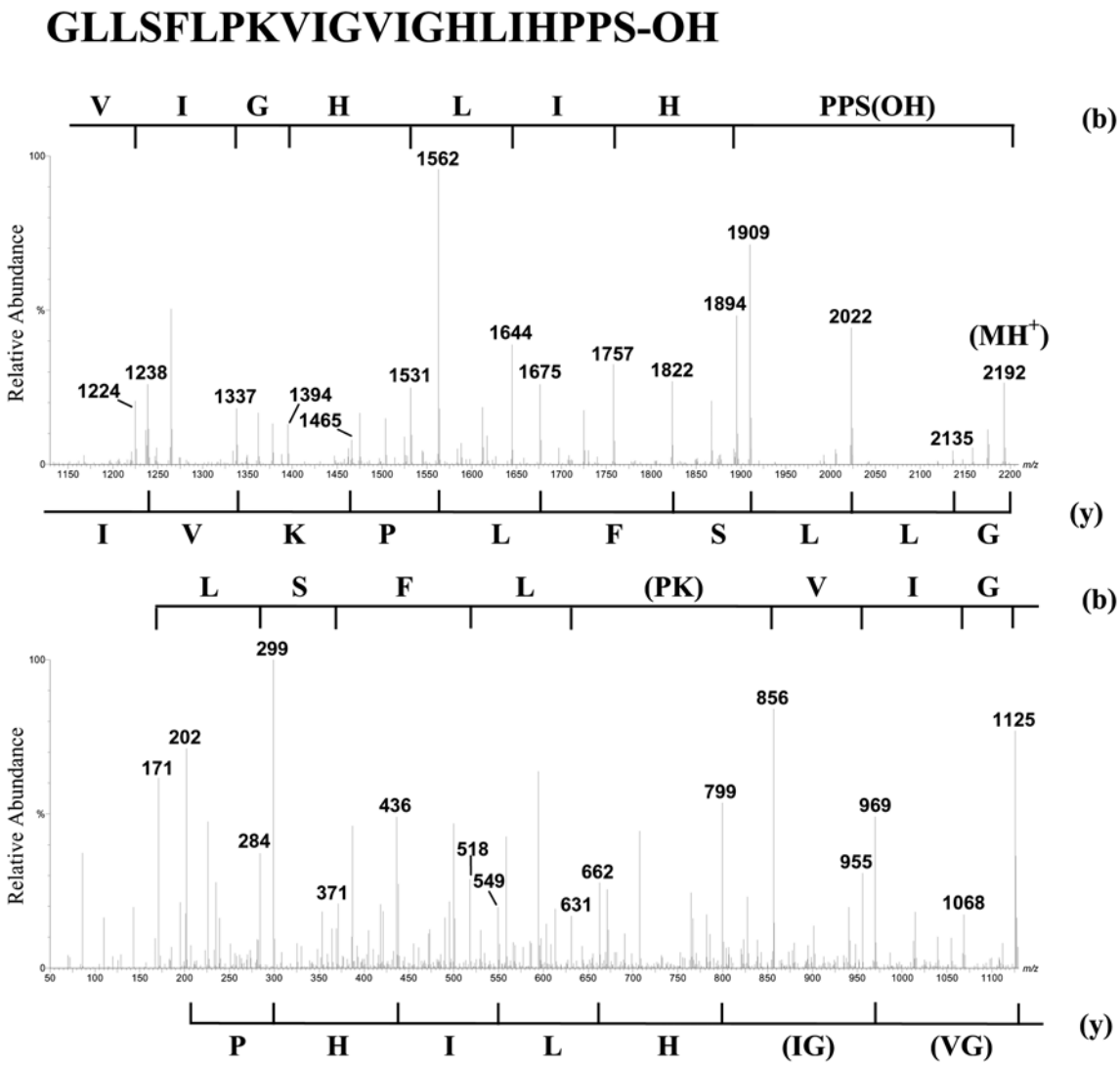


Figure 4.7: Collision-induced MS/MS data for fallaxidin 3.1 illustrating positive ion fragmentations of the $(MH)^+$ species. **b** and **y** cleavage ions are illustrated above and below the spectrum respectively. Magnification ranges: *m/z* 150 – 290 (x3), 350 – 1130 (x3), 1200 – 2180 (x2).

Figure 4.8 and Table 4.4 show MS/MS data from the $[(M-H)^- - 2CH_2O]^-$ ion of fallaxidin 3.1. The loss of two CH_2O groups from the $(M-H)^-$ parent ion indicates the presence of two Ser residues. The α cleavages shown above the spectrum and the β` cleavages below provide sequencing information. In particular, the negative ion spectrum identifies the last two residues Pro and Ser. A combination of the positive and negative ion data provides the full sequence of fallaxidin 3.1, apart from differentiating between Leu and Ile residues.

The *m/z* values of the MH^+ and $(M-H)^-$ parent ions of fallaxidin 3.1 indicate the presence of a C-terminal free acid group. However, the MH^+ ion loses both 17 and 18 Da which is

ambiguous. As a consequence, two peptides corresponding to the fallaxidin 3.1 sequence, one with a C-terminal amide and the other with a free acid were synthesised commercially. Comparison of the MS data from the native fallaxidin 3.1 and the two synthetic peptides identified that fallaxidin 3.1 has a C-terminal free acid. Additionally, the absence of a C-terminal amide donating Gly in the cDNA sequence after the fallaxidin 3.1 peptide sequence, confirmed a free acid terminal (Section 7.2.3).

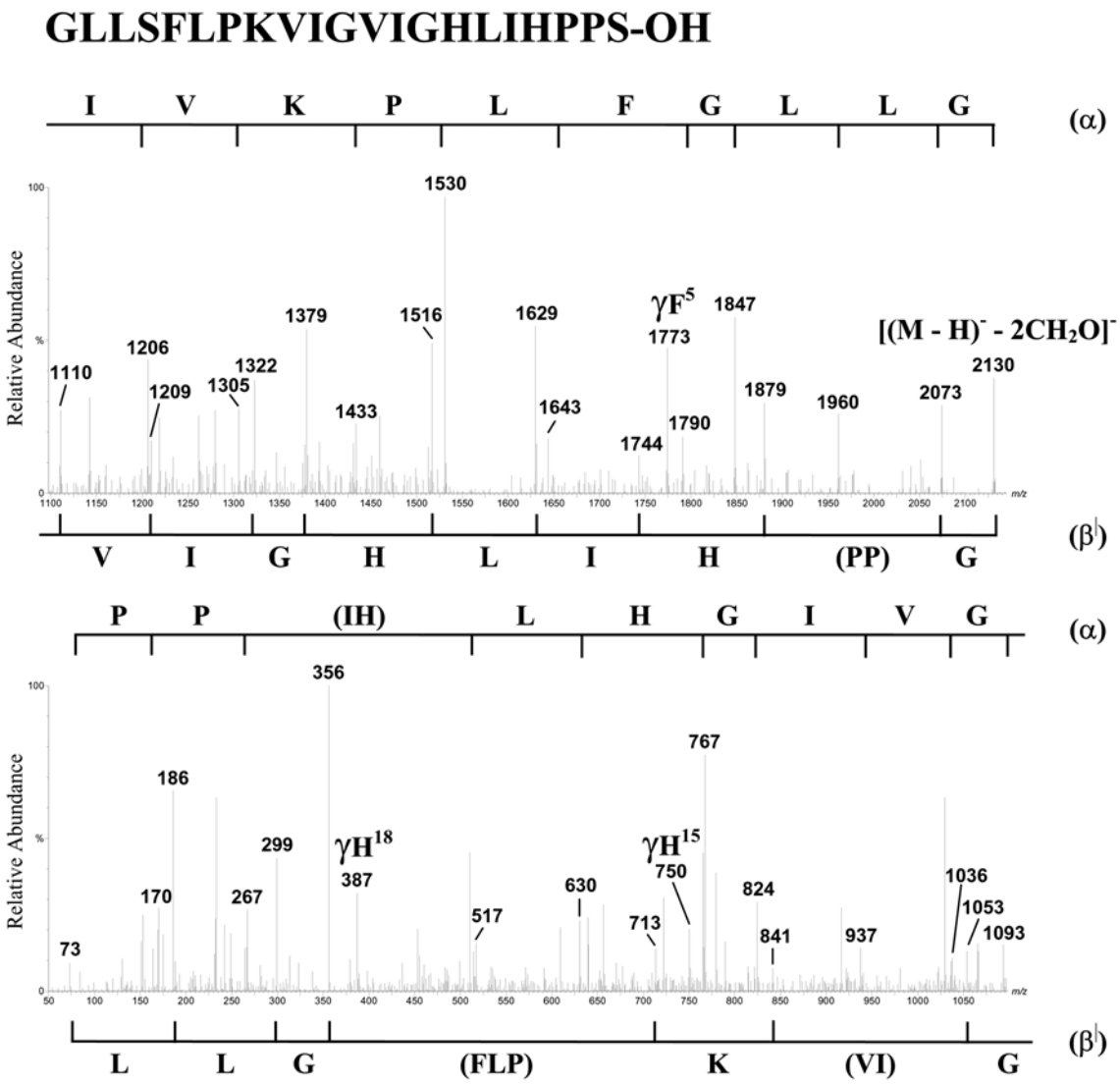


Figure 4.8: Collision-induced MS/MS data for fallaxidin 3.1 illustrating negative ion fragmentations of the $[(M-H)^- - 2CH_2O]^-$ species. α and β` cleavage ions are illustrated above and below the spectrum respectively. Magnification ranges: *m/z* 50 – 325 (x4), 375 – 1100 (x4), 1100 – 1525 (x6), 1600 – 2100 (x6).

Table 4.4: Mass spectral data for fallaxidin 3.1

Fallaxidin 3.1 GLLSFLPKVIGVIGHLIHPPS-OH	
$(MH)^+$ *m/z* 2192	
b ions	*m/z* 1894, 1757, 1644, 1531, 1394, 1337, 1224, 1125, 1068, 955, 856, 631, 518, 371, 284, 177 [LSFL(PK)VIGVIGHLIH(PPS)-OH]
y ions	*m/z* 2135, 2022, 1909, 1822, 1675, 1562, 1465, 1337, 1238, 1125, 969, 799, 662, 549, 436, 299, 202 [GLLSFLPKVI(GV)(IG)HLIHP]
Sequence: GLLSFLPKVIGVIGHLIHPPS-OH	
$(M-H)^-$ *m/z* 2190	
$[(M-H)^- - 2CH_2O]^-$ *m/z* 2130	
α ions from $[(M-H)^- - 2CH_2O]^-$	*m/z* 2073, 1960, 1847, 1790, 1643, 1530, 1433, 1305, 1206, 1093, 1306, 937, 824, 767, 630, 517, 267, 170, 73 [GLLGFLPKVIGVIGHL(IH)PP]
β` ions from $[(M-H)^- - 2CH_2O]^-$	*m/z* 2073, 1879, 1744, 1629, 1516, 1379, 1322, 1209, 1110, 1053, 841, 713, 356, 299, 186, 73 [LLG(FLP)K(VI)GVIGHLIH(PP)G-OH]
γF^5 *m/z* 1774, γH^{15} *m/z* 750, γH^{18} *m/z* 337	
Sequence: GLLSFLPKVIGVIGHLIHPPS-OH	

4.2.3 Secondary Structure Determination of Fallaxidin 3.1

4.2.3.1 NMR Assignment

NMR spectra for the native fallaxidin 3.1 and its amide modification were recorded in the membrane mimicking solvent TFE/H_2O (1:1 v/v) and in DPC micelles. No differences were observed between the overall structures of the native fallaxidin 3.1 (Appendix B) and its amide, thus the solution structure of the more biologically active amide modification is determined here and for simplicity, referred to as fallaxidin 3.1(NH_2).

NMR spectra for the fallaxidin 3.1(NH_2) in DPC micelles were acquired at pH 6.0 to simulate physiological conditions, while in aqueous TFE the spectra were acquired at pH 2.1. The only ionisable group between pH 2 and 6 is the C-terminal amide group, thus the structures at pH 6 should be comparable to those formed at pH 2.1 in aqueous TFE.

The proton resonances were assigned from TOCSY, COSY and NOESY spectra using the sequential assignment procedure outlined by Wüthrich [174]. Broader peaks were observed for fallaxidin 3.1(NH_2) in DPC micelles as a result of a larger correlation time, due to a slower tumbling rate of the peptide-micelle complexes relative to the peptide alone. Despite this, the peaks are still sufficiently sharp for high-resolution NMR studies to be carried out.

There were several Pro residues (Pro7, Pro19 and Pro20) present in the peptide, however no additional series of peaks were observed in the TOCSY or NOESY spectra of either solvent system. This indicates that only one conformer of the Pro residues is present in the secondary structure of fallaxidin 3.1(NH_2).

In aqueous TFE, there was overlap observed in the cross sections for the residues. Leu3 and Ser21 as well as Phe5 and Val9 were found to have coincidental amide chemical shifts. In DPC micelles almost half of the residues had amide proton resonances within a narrow range (7.74 – 7.98 ppm), thus creating a considerable amount of ambiguity in assignment. Differentiation between the resonances in the individual spin systems was achieved through the use of COSY spectra and NOE connectivities. Individual spin systems can be seen in the partial TOCSY and NOESY spectra for aqueous TFE (Figure 4.9) and DPC

micelles (Figure 4.10). The d_{NN} NOE cross-peaks enabled assignment and differentiation of the coincidental residues. The high number of d_{NN} NOE connectivities observed in the NOESY spectrum allowed the sequential assignment of the majority of the residues.

The αC resonances were assigned using the ^{13}C-^{1}H HSQC spectrum, by observation of the αH ^{1}H/αC ^{13}C region of the spectra (Figure 4.11 and Figure 4.12). Assignment of all the αC resonances except for Pro19 was possible from the HSQC spectrum in aqueous TFE, however due to the low signal-to-noise ratio for the DPC micelle sample, not all αC assignments could be made in this solvent system.

Table 4.5 and Table 4.6 present a summary of all the ^{1}H and ^{13}C chemical shifts for fallaxidin 3.1(NH_2) in aqueous TFE and DPC micelles respectively.

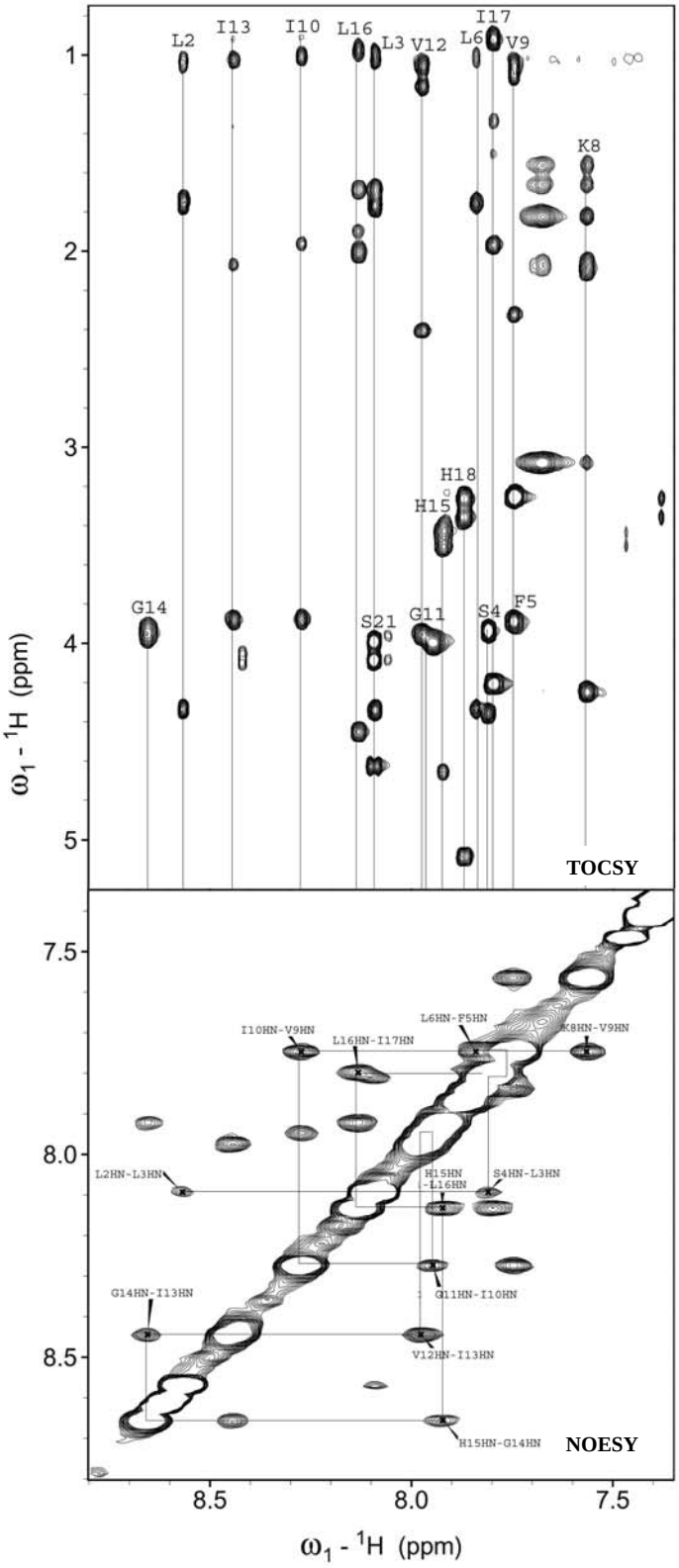


Figure 4.9: Partial TOCSY and NOESY spectra of fallaxidin 3.1(NH_2) in TFE/H_2O (1:1 v/v). In the TOCSY spectrum, vertical lines connect resonances in the same spin system. NOEs between sequential amide protons are indicated in the NOESY spectrum.

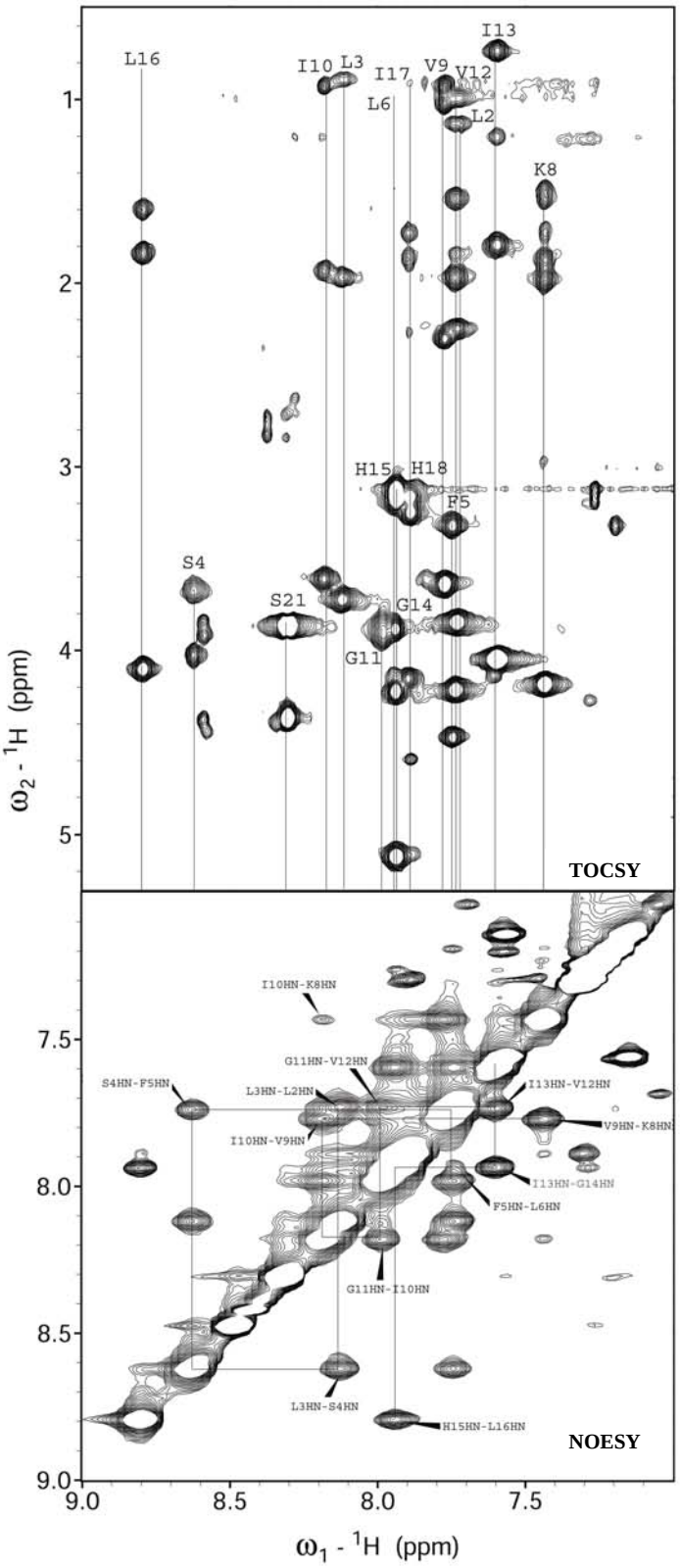


Figure 4.10: Partial TOCSY and NOESY spectra of fallaxidin 3.1(NH_2) in DPC micelles. In the TOCSY spectrum, vertical lines connect resonances in the same spin system. NOEs between sequential amide protons are indicated in the NOESY spectrum.

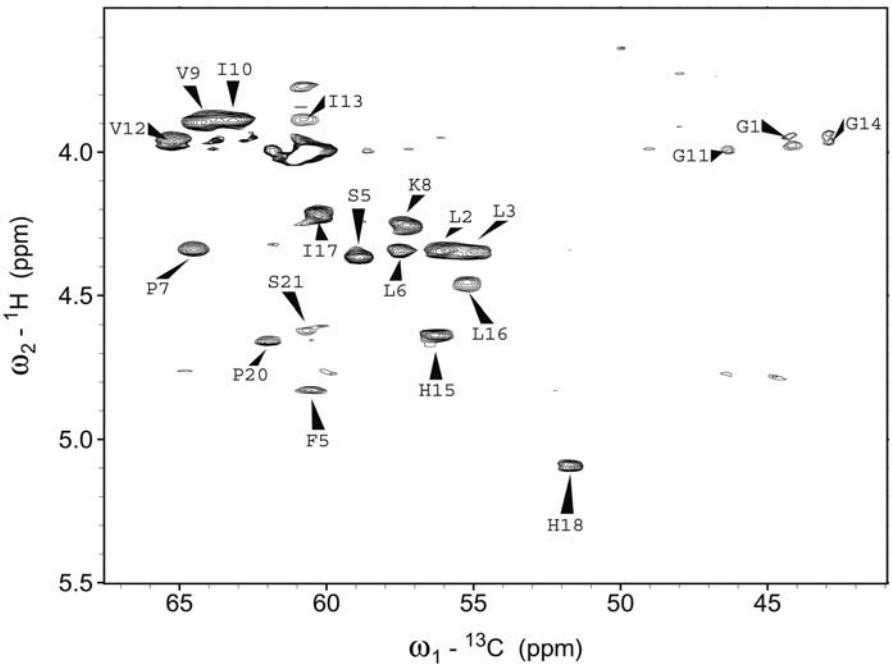


Figure 4.11: Partial ^{13}C-^{1}H HSQC spectrum of fallaxidin 3.1(NH_2) in TFE/H_2O (1:1 v/v). Only the αH ^{1}H/αC ^{13}C connectivities for the residues are indicated.

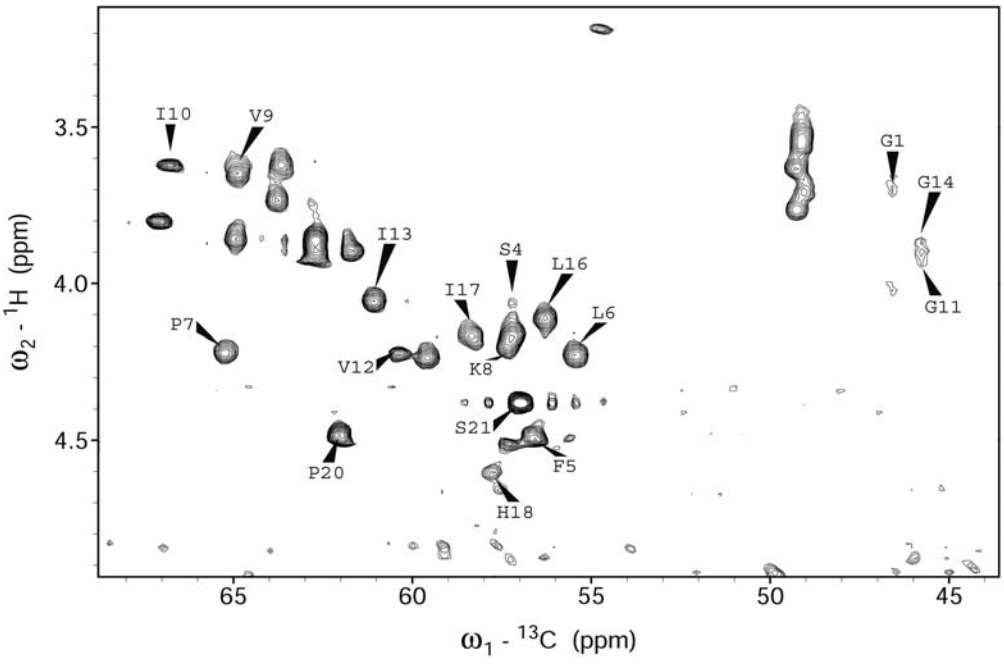


Figure 4.12: Partial ^{13}C-^{1}H HSQC spectrum of fallaxidin 3.1(NH_2) in DPC micelles. Only the αH ^{1}H/αC ^{13}C connectivities for the residues are indicated.

Table 4.5: ^{1}H and ^{13}C chemical shifts for fallaxidin 3.1(NH_2) in TFE/H_2O (1:1 v/v), pH 2.10, 25 ºC. n.o. indicates that the resonance was not observed.

Residue	Chemical Shift				
	HN	**αH**	**βH**	**Other H**	**α^{13}C**
Gly1	n.o.	4.05, 3.95	-	-	44.26
Leu2	8.57	4.34	1.75	γ-CH n.o	56.03
				δ-CH_3 1.04	
Leu3	8.09	4.34	1.77	γ-CH 1.69	54.97
				δ-CH_3 1.04, 0.98	
Ser4	7.81	4.36	3.94	-	58.91
Phe5	7.75	4.82	3.26	H 2,6 7.31	60.53
				H 3,5 7.40	
				H 4 7.36	
Leu6	7.84	4.34	1.76	γ-CH n.o	57.48
				δ-CH_3 1.01	
Pro7	-	4.33	2.44, 1.88	γ-CH_2 2.20, 2.02	64.53
				δ-CH_2 3.79, 3.60	
Lys8	7.57	4.25	2.08	γ-CH_2 1.66, 1.56	57.32
				δ-CH_2 1.83	
				ε-CH_2 3.08	
				ε-NH_3^+ 7.68	
Val9	7.75	3.89	2.32	γ-CH_3 1.11, 1.05	64.17
Ile10	8.27	3.88	1.96	γ-CH_2 1.67, 1.19	63.18
				γ-CH_3 n.o.	
				δ-CH_3 0.91	
Gly11	7.95	4.00	-	-	46.34
Val12	7.98	3.96	2.41	γ-CH_3 1.16, 1.06	65.25
Ile13	8.44	3.88	2.07	γ-CH_2 1.77, 1.67	60.78
				γ-CH_3 1.36	
				δ-CH_3 1.03	
Gly14	8.66	3.97, 3.93	-	-	42.93
His15	7.92	4.66	3.51, 3.43	H 2 7.47	56.25
				H 4 8.60	

Table 4.5: continues.

Residue	Chemical Shift				
	HN	**αH**	**βH**	**Other H**	**α^{13}C**
Leu16	8.13	4.45	2.01, 1.90	γ-CH 1.69	55.22
				δ-CH_3 0.98, 0.96	
Ile17	7.80	4.21	1.98	γ-CH_2 1.51, 1.34	60.28
				γ-CH_3 n.o.	
				δ-CH_3 0.92	
His18	7.87	5.08	3.36, 3.26	H 2 7.38	51.70
				H 4 8.62	
Pro19	-	4.82	2.49, 1.83	γ-CH_2 2.10	n.o.
				δ-CH_2 3.82, 3.61	
Pro20	-	4.65	2.40, 1.98	γ-CH_2 2.15	61.93
				δ-CH_2 3.89, 3.75	
Ser21	8.09	4.63	4.09, 3.99	-	60.63

Table 4.6: ^{1}H and ^{13}C chemical shifts for fallaxidin 3.1(NH_2) in DPC micelles (40 molar equivalents), pH 6.00, 25 ºC. n.o. indicates that the resonance was not observed.

Residue	Chemical Shift				
	HN	**αH**	**βH**	**Other H**	**$\alpha^{13}C$**
Gly1	n.o.	3.71, 3.61	-	-	45.58
Leu2	7.74	3.84	2.25	γ-CH 1.54	n.o.
				δ-CH_3 1.13	
Leu3	8.12	3.72	1.97, 1.77	γ-CH 1.15	n.o.
				δ-CH_3 0.89, 0.79	
Ser4	8.62	4.03	3.68	-	57.31
Phe5	7.75	4.74	3.32	H 2,6 7.20	56.61
				H 3,5 7.57	
				H 4 7.14	
Leu6	7.94	4.22	1.98, 1.87	γ-CH 1.55	55.42
				δ-CH_3 0.90	
Pro7	-	4.20	2.38, 1.90	γ-CH_2 2.18	65.22
				δ-CH_2 3.69	
Lys8	7.44	4.19	1.98, 1.87	γ-CH_2 1.54, 1.49	57.26
				δ-CH_2 1.72	
				ε-CH_2 2.97	
				ε-NH_3^+ 7.90	
Val9	7.77	3.64	2.30	γ-CH_3 1.02, 0.93	64.89
Ile10	8.18	3.61	1.93	γ-CH_2 1.64, 1.20	66.76
				γ-CH_3 0.93	
				δ-CH_3 0.83	
Gly11	7.98	3.92, 3.87	-	-	45.78
Val12	7.74	4.22	1.97	γ-CH_3 1.00, 0.90	60.36
Ile13	7.60	4.05	1.79	γ-CH_2 1.54, 1.32	61.09
				γ-CH_3 1.20	
				δ-CH_3 0.74	
Gly14	7.94	3.89	-	-	45.77
His15	7.94	5.12	3.18, 3.11	H 2 7.27	51.38
				H 4 8.48	

Table 4.6: continues.

Residue	Chemical Shift				
	HN	**αH**	**βH**	**Other H**	**α^{13}C**
Leu16	8.80	4.10	1.84	γ-CH 1.60	56.28
				δ-CH_3 0.98, 0.90	
Ile17	7.90	4.15	1.87	γ-CH_2 1.73, 1.63	58.35
				γ-CH_3 1.01	
				δ-CH_3 0.90	
His18	7.89	4.59	3.25, 3.16	H 2 7.30	57.76
				H 4 7.43	
Pro19	-	4.74	2.33, 1.91	γ-CH_2 2.02	n.o.
				δ-CH_2 3.53	
Pro20	-	4.46	2.30, 1.96	γ-CH_2 2.03	62.04
				δ-CH_2 3.76, 3.62	
Ser21	8.31	4.37	3.88, 3.84	-	57.00

4.2.3.2 Secondary Shifts

The variations in chemical shifts for the amino acids in the peptide from their corresponding random coil values (as determined in water [190]) are referred to as the Δδ of the residues. The Δδ for αH ^{1}H and αC ^{13}C are smoothed over n = ± 2 residues, while the amide ^{1}H Δδ remain unsmoothed. The amino acid sequence was plotted against the smoothed Δδ value for αH ^{1}H, αC ^{13}C and the unsmoothed Δδ for amide ^{1}H resonances (Figure 4.13). As discussed in Section 3.6, upfield shifts in resonances are illustrated by negative Δδ values, whilst positive values indicate a downfield shift in the resonances.

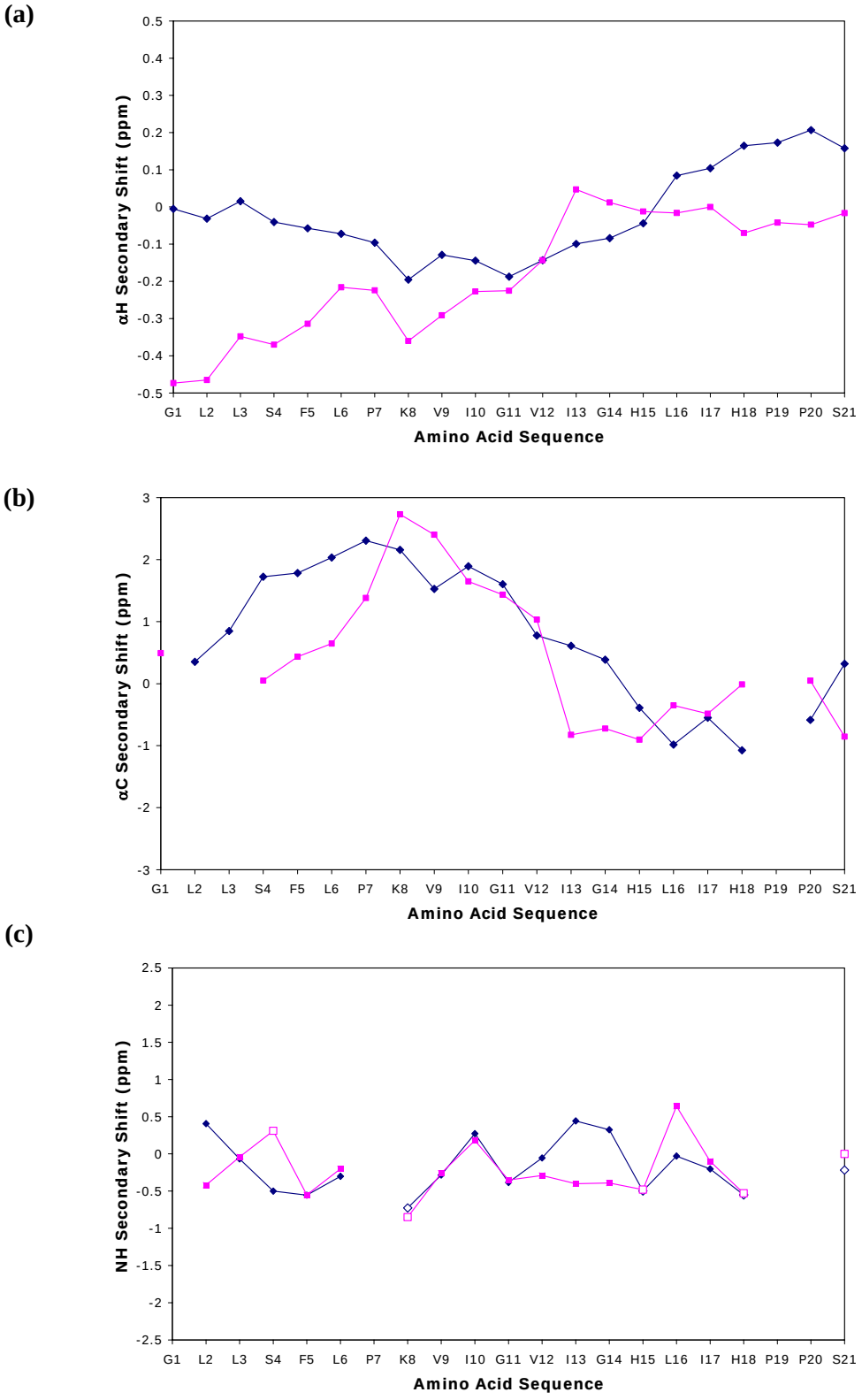


Figure 4.13: (a) αH ^{1}H and **(b)** αC ^{13}C secondary shifts of fallaxidin 3.1(NH_2) in TFE/H_2O (1:1 v/v) (blue) and DPC micelles (pink), smoothed over n = ± 2 residues. **(c)** NH ^{1}H secondary shifts of fallaxidin 3.1(NH_2). Negative values indicate an upfield shift from the random coil values and positive values indicate a downfield shift. Open symbols indicate hydrophilic residues.

The αH ^{1}H Δδ plot illustrates an upfield shift for residues Phe5 to Ile13, which is consistent with a region of helical structure in aqueous TFE. However, αH chemical shifts are typically shifted upfield by approximately 0.39 ppm in a helical structure [193, 247]. The magnitude of the αH shifts here (~ 0.2 ppm) suggests a more flexible helix-like structure may be present. In contrast, residues Leu16 to Ser21 display downfield shifts consistent with β sheet or extended conformers in this region. However, once again, the magnitude of the shifts (0.2 ppm) suggests a degree of flexibility (c.f. ~ 0.37 ppm shifts in β sheet [193]). Some flexibility is expected at the ends of the peptide and this is illustrated by the near zero values at the N-terminal. These conclusions are supported by the αC Δδ.

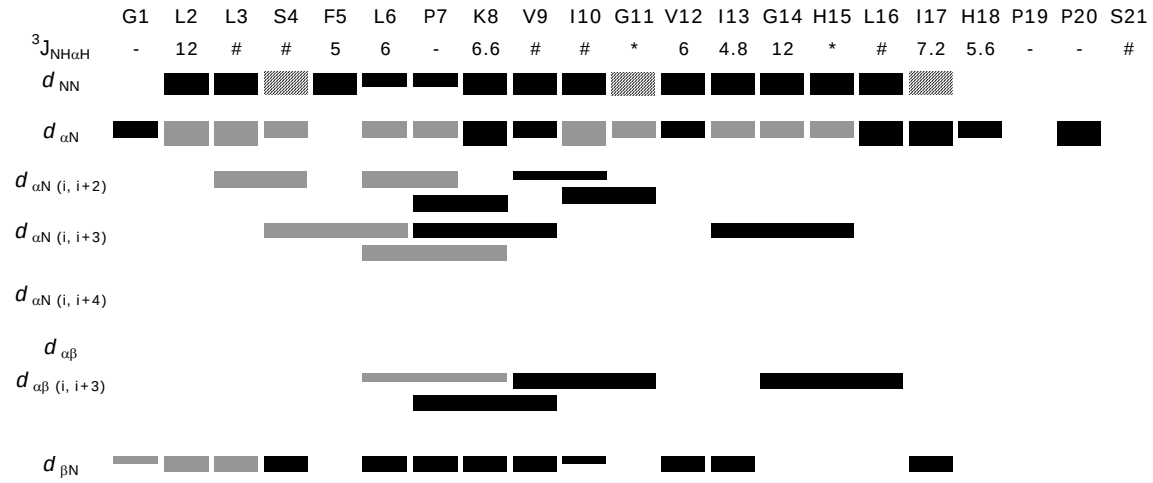
Similarly, in DPC micelles, the fallaxidin 3.1(NH_2) αH ^{1}H Δδ plot illustrates an upfield shift for residues in the N-terminus with some disruption to the helical shifts around Ile13. As expected in the C-terminal region consisting of two sequential Pro resides, flexibility in the secondary structure is implied by near zero values. This conclusion is also reflected in the αC ^{13}C Δδ plot.

The amide proton Δδ values show an upfield shift for the hydrophilic residues relative to those of the hydrophobic residues, which display a downfield shift. A periodic variation over three to four residues is present along the length of the peptide with the exception of the residues neighbouring Ile13 in DPC micelles. These characteristics are consistent with a helical region displaying amphipathic nature [197, 198].

4.2.3.3 NOE Connectivities

Figure 4.14 illustrates a summary of the diagnostic NOEs present in the NOESY spectrum of fallaxidin 3.1(NH_2) in both solvent environments. In aqueous TFE, the absence of medium range NOEs at the C-terminus of the peptide suggests that an extended or flexible conformation may be present in this region. The NOE connectivity pattern for residues Phe5 to Val12 implies that this peptide region may contain some helical structure. In DPC, there appears to be disruptions to defined secondary structure as indicated by the absence of medium range NOE connectivities at the C-terminus from His18 and in the regions of Ser4 to Leu6 and Gly11 to Ile13. The remaining regions of the peptide have NOE connectivity patterns consistent with helical secondary structures.

(a)



(b)

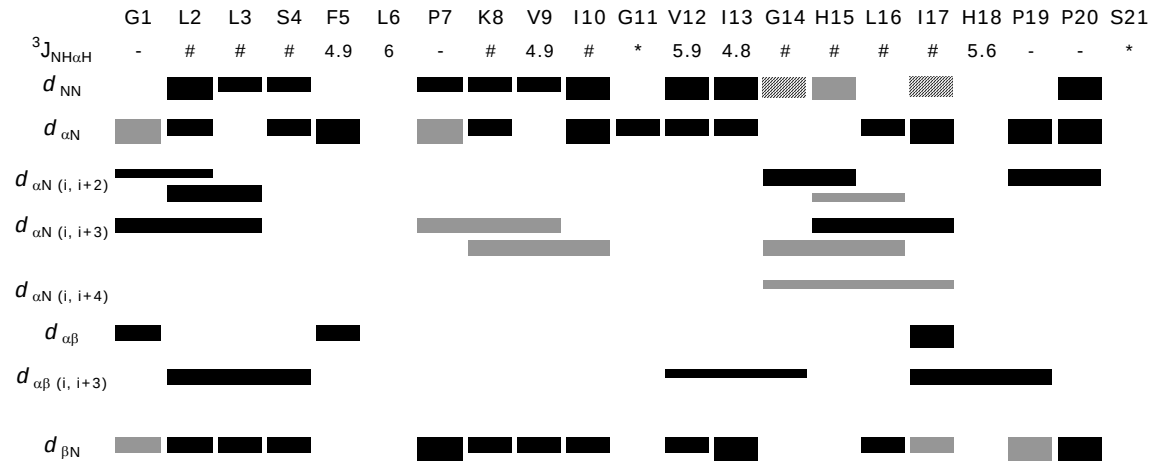


Figure 4.14: A summary of the diagnostic NOEs used in the structure calculations for fallaxidin 3.1(NH_2) in **(a)** TFE/H_2O (1:1 v/v) and **(b)** DPC micelles. The thickness of the bar indicates the relative strength of the NOE (strong < 3.1 Å, medium 3.1 – 3.7 Å, weak > 3.7 Å). Grey shaded bars represent ambiguous NOEs. For Pro where no amide proton is present, NOEs to δH are shown. $^3J_{NH\alpha H}$ values are indicated where applicable. # indicates coupling constants that were not resolved due to signal overlap, * indicates no coupling constant was detected.

Due to the presence of two Pro residues at the C-terminus, defined secondary structure in this region is expected to be disrupted. The extent of *cis-trans* isomerization can be determined from the observed NOE connectivities about the Pro imide bond [248]. As $d_{\alpha\delta}$ and not $d_{\alpha\alpha}$ connectivities were present in both solvent conditions, it is likely that the *trans* isomer is the dominant isomer for both Pro19 and Pro20. Likewise, this was also observed for Pro7.

4.2.3.4 Structure Calculations

The NOESY spectrum for fallaxidin 3.1(NH_2) was fully assigned and the cross-peak values quantified and converted to distance restraints in accordance with the method of Xu *et al*. [208]. A total of 229 and 275 non-redundant restraints were generated for aqueous TFE and DPC micelle solutions respectively, and are summarised in Table 4.7. Of these distance restraints, only a relatively small number are ambiguous.

Table 4.7: The experimental distance restraints obtained from the NOESY spectrum of fallaxidin 3.1(NH_2).

	No. of Experimental Restraints	
	TFE/H_2O	**DPC**
Sequential NOEs	45	71
Medium-range NOEs	12	30
Long-range NOEs	-	-
Intra-residue NOEs	125	130
Ambiguous NOEs	37	44
Total	229	275

The amide region of the high-resolution 1D ^1H spectrum was used to calculate the $^3J_{NH\alpha H}$ coupling constants (Figure 4.14). A degree of overlap was observed in this region of the 1D ^1H spectrum for both solvent systems and as a result, many coupling constants were unable to be determined. Attempts were made to determine coupling constants from the COSY spectra using the method of Kim and Prestegard [203], however the line widths were too broad to determine these values accurately. In aqueous TFE the values for $^3J_{NH\alpha H}$ coupling constants are consistent with helical secondary structure for residues Phe5 to Ile13, whilst those for residues Gly14 and Ile17 suggest that the C-terminal of the peptide displays little defined structure. In DPC micelles, all measurable $^3J_{NH\alpha H}$ coupling constants are consistent with regions of helical secondary structure. This supports the conclusions from the $\Delta\delta$ plots (Figure 4.13) and NOE connectivity patterns (Figure 4.14).

Initially, sixty structures were generated using random backbone torsion angles. The structures were subsequently subjected to RMD and SA calculations and the resultant twenty lowest potential energy structures were chosen for analysis. The lowest energy

structures for fallaxidin 3.1(NH_2) in aqueous TFE and DPC micelles are illustrated in Figure 4.15 and Figure 4.16 respectively. The backbone atoms are superimposed over well-defined residues and the entire backbone in both solvent environments.

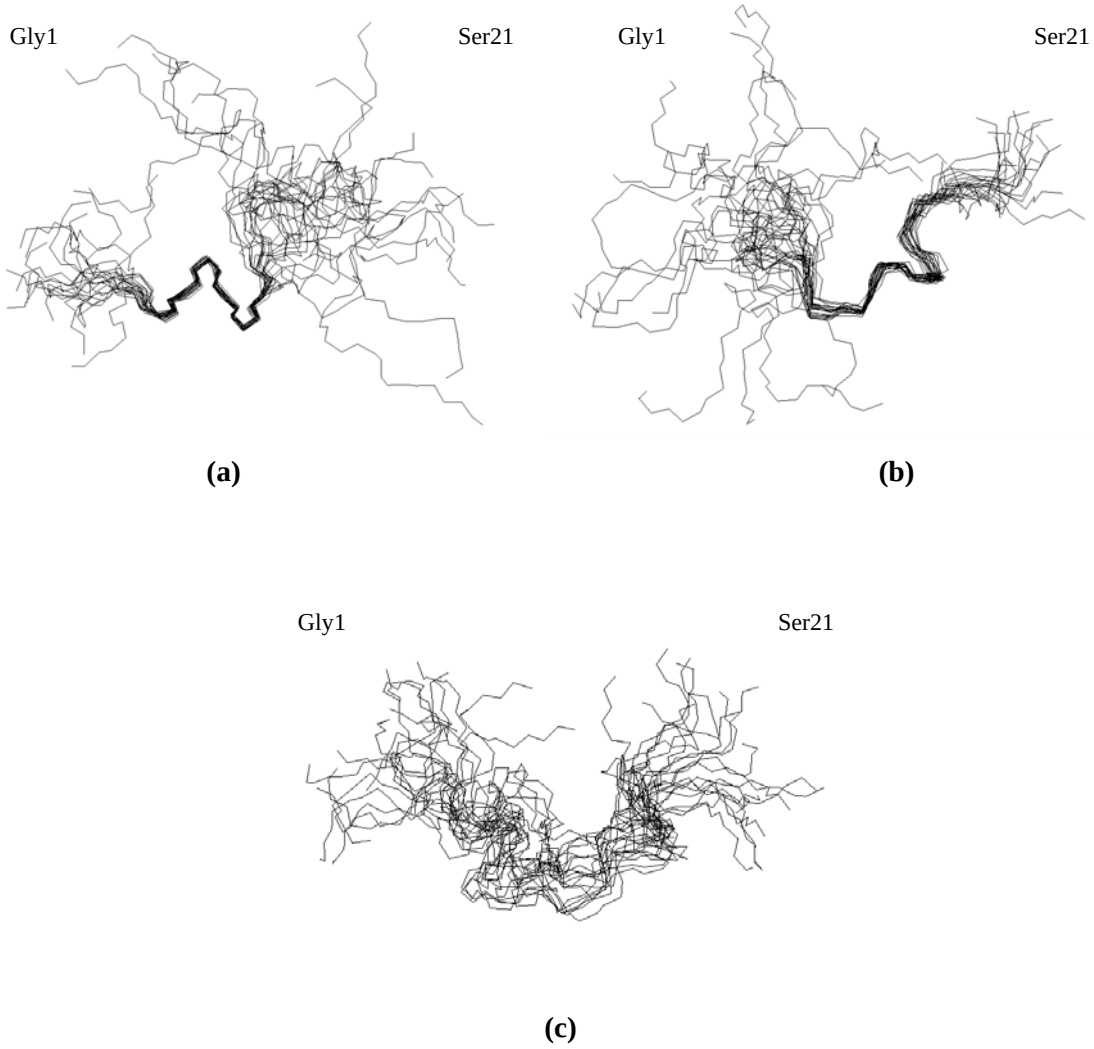


Figure 4.15: The twenty most stable calculated structures of fallaxidin 3.1(NH_2) in TFE/H_2O (1:1 v/v). The structures have been superimposed over the backbone atoms of the well-defined **(a)** residues 5 – 10 and **(b)** residues 12 – 19 and **(c)** superimposed over the entire backbone.

In aqueous TFE (Figure 4.15), there appears to be some α helical secondary structure for residues 5 to 10 and a combination of α helical and β strand structure in the C-terminal region. This can be seen clearly in Figure 4.17. The defined secondary structure is clearly disrupted by the presence of Gly and Pro residues at intervals of three to four residues. In DPC micelles (Figure 4.16), fallaxidin 3.1(NH_2) displays more defined regions of secondary structure resembling an α helix. However, structure disruptions by Gly11 and Gly14 result in the central region of the peptide between these two residues displaying a

more extended strand like structure. Overall, the secondary structure of fallaxidin 3.1(NH_2) appears to be helical about its length with a disruption to this structure in the region of Gly11 and Gly14.

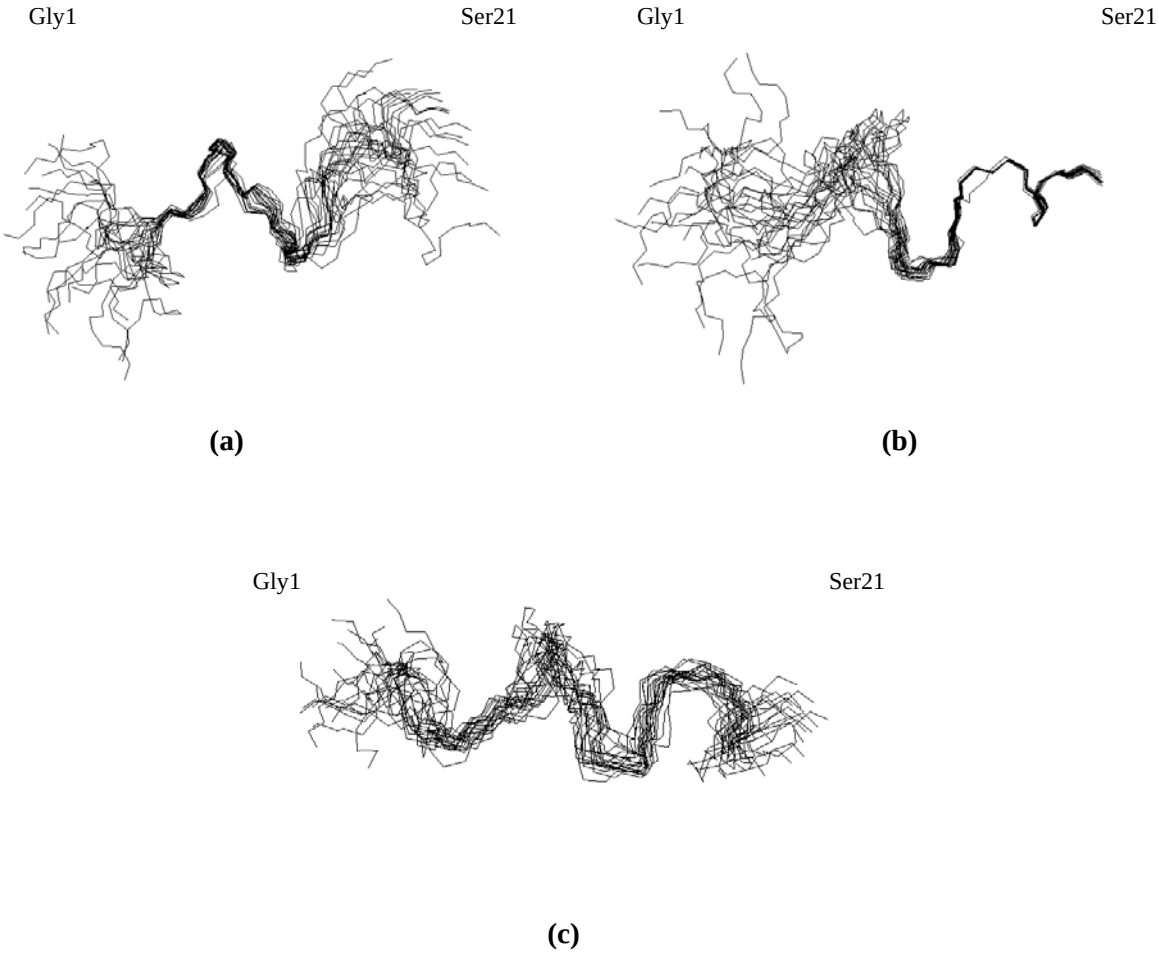


Figure 4.16: The twenty most stable calculated structures of fallaxidin 3.1(NH_2) in DPC micelles. The structures have been superimposed over the backbone atoms of the well-defined **(a)** residues 7 – 13 and **(b)** residues 15 – 20 and **(c)** superimposed over the entire backbone**.**

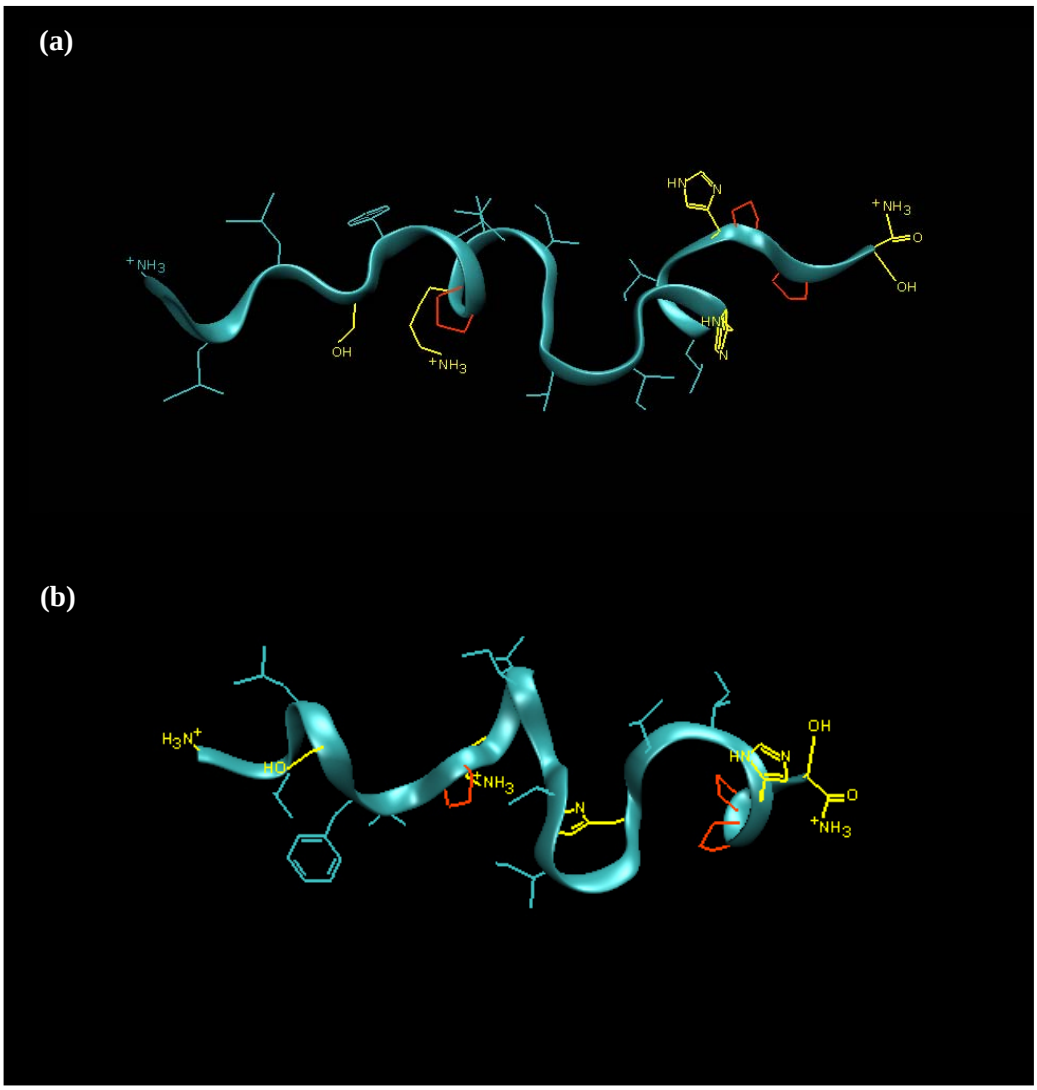


Figure 4.17: The lowest calculated potential energy structure of fallaxidin 3.1(NH_2) in **(a)** TFE/H_2O (1:1 v/v) and **(b)** DPC micelles. Yellow indicates residues that are hydrophilic in nature. Pro residues are shown in red.

The twenty lowest energy structures illustrate high convergence over residues 5 – 10 and 12 – 19 in aqueous TFE and 7 – 13 and 15 – 20 in DPC micelles. The RMSD geometry values of backbone atoms are relatively low, indicating little deviation from the mean geometry. The RMSD values (Table 4.8) for all atoms are considerably larger than previously reported structures (c.f caerin 1.1 in DPC, all backbone atoms – 0.64 ± 0.22 Å [249]), possibly highlighting that there are regions of conformational flexibility within the secondary structure. This is particularly evident for the aqueous TFE system.

The final structures had only minor deviations from idealised covalent geometry (≤ 0.05 Å for bonds and ≤ 5° for angles and impropers). In aqueous TFE and DPC micelle

samples, there were 4 and 8 violations greater than 0.3 Å respectively. The maximum violations were 0.56 and 0.52 Å respectively.

Table 4.8: The structural statistics for the twenty lowest energy structures of fallaxidin 3.1(NH_2) following RMD/SA calculations.

Energies (kcal mol^{-1})	**TFE/H_2O**	**DPC**
E_{total}	18.046 ± 0.492	25.470 ± 0.912
E_{bond}	0.537 ± 0.029	1.140 ± 0.097
E_{angle}	8.143 ± 0.206	10.823 ± 0.390
$E_{improper}$	0.242 ± 0.033	0.335 ± 0.049
E_{vdm}	9.097 ± 0.391	13.060 ± 0.695
E_{NOE}	0.023 ± 0.020	0.111 ± 0.078
E_{cdih}	0.005 ± 0.006	0.001 ± 0.003
Well defined residues	5-10, 12-19	2-3, 7-13, 15-20

RMSD from mean geometry (Å)				
Heavy atoms of well-defined residues	(5-10)	1.104 ± 0.244	(7-13)	0.808 ± 0.308
Backbone atoms of well-defined residues	(5-10)	0.268 ± 0.008	(7-13)	0.440 ± 0.224
Heavy atoms of well-defined residues	(12-19)	1.160 ± 0.278	(15-20)	0.449 ± 0.232
Backbone atoms of well-defined residues	(12-19)	0.660 ± 0.206	(15-20)	0.139 ± 0.117
All heavy atoms		4.289 ± 0.721		2.263 ± 0.466
All backbone atoms		3.678 ± 0.762		1.922 ± 0.500

Analysis of the AOP (S, ϕ and ψ) of the twenty lowest energy structures illustrated that for the aqueous TFE and DPC micelle samples of fallaxidin 3.1(NH_2), fourteen and fifteen of the twenty-one residues were well-defined (S > 0.9) respectively. Little variation is expected in regions of well-defined residues. As expected, Gly11 and Gly14 (with the exception of Gly14 in aqueous TFE) do not display well-defined nature, which is consistent with disruptions in the secondary structure of the peptide. In DPC, all residues (excluding Pro and Gly residues) of fallaxidin 3.1(NH_2) displayed AOP consistent with helical structure about its length with obvious disruptions created by Pro and Gly residues. The Ramachandran plot (Figure 4.18) for the average ϕ and ψ values illustrate that all of

the residues in the peptide fall into the favourable or allowable regions, further supporting the quality of the calculated structures in representing the solution structure of fallaxidin 3.1(NH_2).

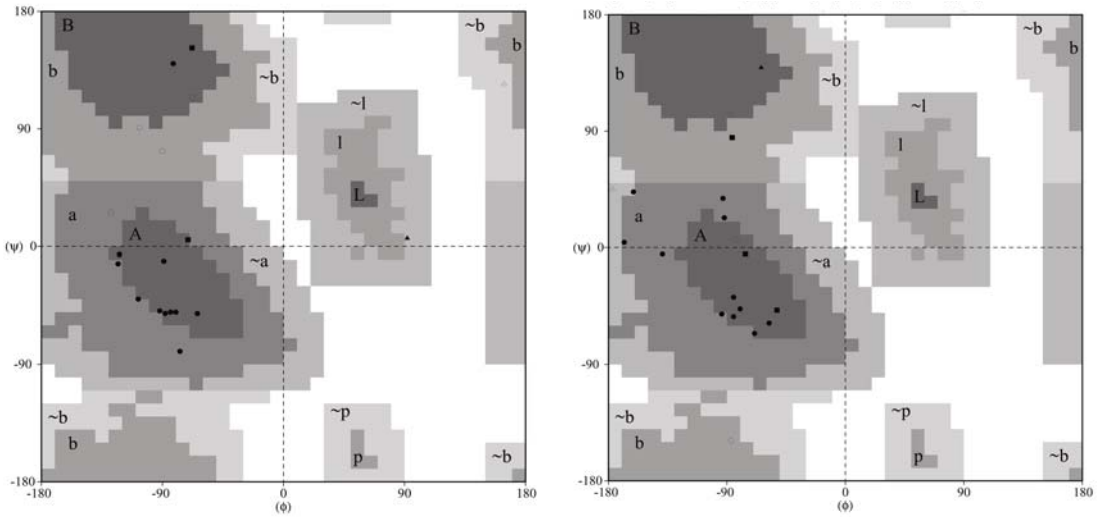


Figure 4.18: Ramachandran plot for the well-defined residues of fallaxidin 3.1(NH_2) in **(a)** TFE/H_2O (1:1 v/v) and **(b)** DPC micelles. Favourable regions are labelled A and B for α helical or β strand structures respectively, while allowable and generous regions are labelled a and b or ~a and ~b.

4.2.4 Biological Activity

Commercially synthesised versions of the fallaxidin peptides were used for biological testing. These synthetic peptides were shown to be identical to the naturally isolated peptides by MS. The biological activity of the fallaxidin peptides is summarised in Table 4.9. The activities of fallaxidin 1.1, 1.2 and 1.3 are discussed in greater detail in Chapter 6, so will only be briefly summarised in this chapter.

Table 4.9: Biological activity of the peptides isolated from *Litoria fallax***.**

Fallaxidin	Sequence	Biological Activity[1]
1.1	YFPIPI-NH_2	a
1.2	YFPIPF-NH_2	a
1.3	YHPF-NH_2	a
1.4	YPPI-NH_2	a
2.1	GLLDLAKHVIGIASKL-NH_2	b
2.2	GLLFLAKHVIGIASKL-NH_2	b
2.3	GLVDFAKHVIGIASKL-NH_2	a
3.1	GLLSFLPKVIGVIGHLIHPPS-OH	b
3.1 mod.	GLLSFLPKVIGVIGHLIHPPS-NH_2	c, d
3.2	GLFSFLPKVIGVIGPLIHPPS-OH	a
4.1	FWPFM-NH_2	a
4.2	FLPLLASLVGGLL-NH_2	a
4.3	FFRVLAKLGKLAK-OH	a

[1]**Activity: (a)** unknown; **(b)** narrow spectrum antimicrobial; **(c)** wide spectrum antimicrobial; **(d)** nNOS inhibitor.

Fallaxidin 1 peptides are believed to be neuropeptides involved in maintaining the animal's dermal physiological environment and/or form a part of the animal's host defence system, however, the exact biological activity is unknown. Fallaxidin 2.1 and 2.2 display limited antibiotic activity against selected Gram-positive organisms (Table 4.10). Both fallaxidin 2.1 and 2.2 have minimal activity against *Leuconostoc lactis*, with the former also displaying minimal activity against *Enterococcus faecalis*. Fallaxidin 3.1 however, is a moderate antimicrobial against selected Gram-positive bacteria (see Table 4.10). In

addition, fallaxidin 3.1(NH_2) modification displays some nNOS activity with an IC_{50} value of 15.4 μM and a Hill slope of 4.0.

Table 4.10: Antibiotic activities (μg.mL^{-1}) of fallaxidins 2.1, 2.2, 3.1 and 3.1(NH_2) against Gram-positive and *Gram-negative organisms. Where no value is indicated, MIC > 100 μg.mL^{-1} and the peptide is considered to be inactive. (MIC: minimal inhibitory concentration required to kill organisms).

	Fallaxidin			
Organism[1]	**2.1**	**2.2**	**3.1**	**3.1(NH_2)**
Bacillus cereus				25
Enterococcus faecalis	100			100
Leuconostoc lactis	100	100	12	3
Listeria innocua				50
Micrococcus luteus			100	12
Staphylococcus aureus				50
Staphylococcus epidermidis			100	25
Streptococcus uberis			50	12
*Enterobacter clocae**				
*Escherichia coli**				

[1]The microorganisms tested are mammalian pathogens but they represent the types of microbes that could be pathogenic to amphibians.

4.2.5 Isolation and Structural Characterisation of Preprofallaxidins

The isolation of preprofallaxidins will be briefly reported here. For a detailed discussion refer to Section 7.2.3.

Nine different preprofallaxidin cDNAs were cloned from the skin secretion of *L. fallax* and each clone encoded either with a single copy or tandem repeats of the fallaxidin peptides at the C-terminal region of each open-reading frame. These clones included previously characterised fallaxidins 1, 2, and 3 (from MS analysis, Section 4.2.2) and six novel fallaxidins, which were named fallaxidin 1.4, 2.3, 3.2 and 4.1 - 4.3 (Table 4.1). Some clones coded for several distinct peptides, flanked by a typical -Lys-Arg- (-KR-) propeptide convertase processing motif [250] and a short acidic spacer (Figure 4.19).

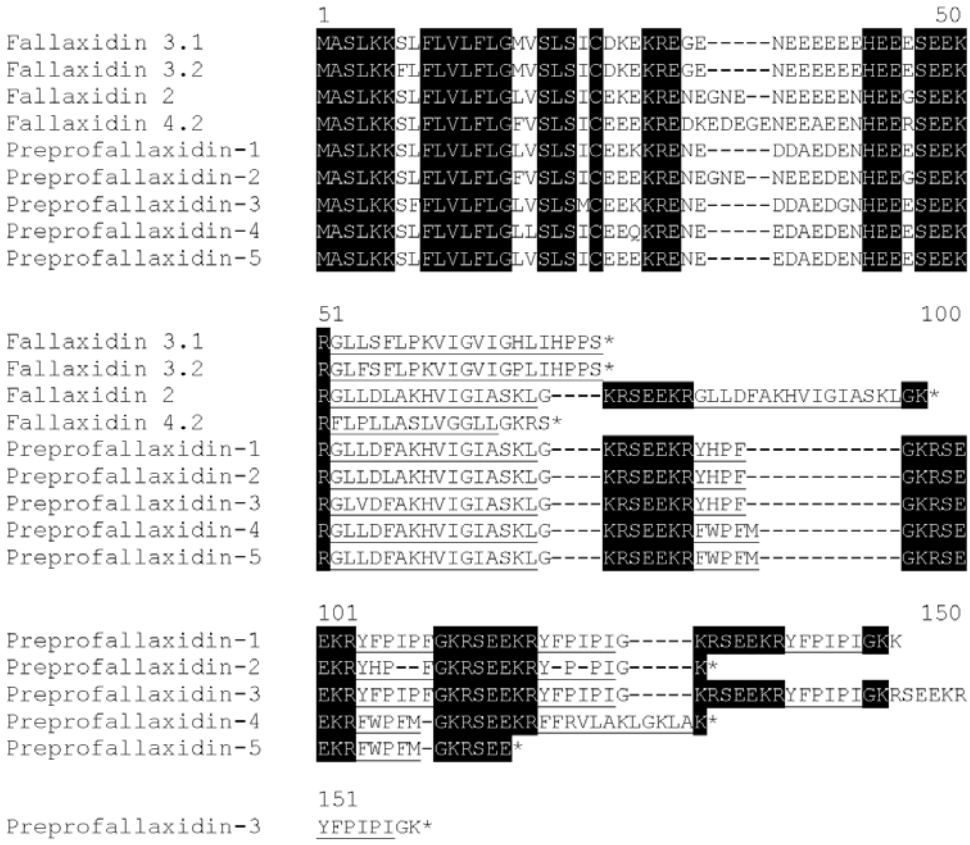


Figure 4.19: Alignment of translated preprofallaxidin open-reading frame of cDNA. Identical amino acid residues in the signal and acidic spacer regions are boxed in black. Fallaxidin peptides are underlined. Stop codons are indicated by *. The high degree of primary sequence conservation of the signal peptides of each preprofallaxidin is evident.

The alignment of preprofallaxidin cDNA revealed that the clones exhibit high degrees of both nucleotide (Figure 7.12) and primary sequence similarities. Putative signal peptide sequences were almost identical (91 % homology, 2 amino acids different) in structure with heterogeneity localised to the acidic spacer peptide and fallaxidin encoded region. The latter regions were flanked N-terminally by a potential propeptide convertase cleavage site of paired basic amino acids (-KR-). The preprofallaxidins exhibit high degrees of structural similarities to other reported peptide precursors from hylids (*L. caerulea* [25], *L. aurea* [251] and *L. infrafrenata* [252]) and ranid frogs [25] with a highly conserved signal peptide (Figure 7.14).

The molecular masses of the novel peptides were deduced from the predicted peptide structures located within the open-reading frames of the respective clones. Of the novel fallaxidin peptides, only two (fallaxidin 4.1 and 4.2) could not be located within the skin secretion fractions as separated by HPLC (Section 4.2.1).

4.3 Discussion

4.3.1 Structure Analysis

NMR spectra and structure analysis indicated no overall difference in the secondary structure of the native fallaxidin 3.1 and its amide modification. As a consequence, the secondary structure of the more active amide modification was investigated. The results obtained from the solution NMR studies indicate that fallaxidin 3.1(NH_2) forms a partial helix at the N-terminal end of the peptide in aqueous TFE. TFE acts as a membrane mimicking solvent by promoting the formation of intramolecular hydrogen bonds [20, 24], therefore it is likely that the structures determined here persist in a bilayer environment. However, in DPC micelles, fallaxidin 3.1(NH_2) displayed a more defined helical structure about its length with a disruption in its centre (Gly11 to Gly14) which varies from that determined in aqueous TFE. It should be noted that from the sequence of fallaxidin 3.1(NH_2) alone, secondary structure prediction programs (Predict Protein: www.predictprotein.org) indicate a degree of helical structure about the length with a high degree of flexibility about the ends.

The differences in structures of fallaxidin 3.1(NH_2) in the two solvent systems may be because micelles are better in stabilising α helices and prevent aggregation of the peptide into β sheet structures [253]. Poor correlations between native protein structures and those determined in aqueous TFE have been observed in β lactoglobulin and α chymotrypsin [254]. This may be a consequence of the ability of TFE to disrupt hydrophobic interactions and denature tertiary structures in some cases [241, 254]. However, fallaxidin 3.1(NH_2) is relatively small and lacks tertiary structure, thus TFE-induced disruptions may not have a significant contribution to the differences in secondary structures. Despite the differences in conformation, the generated structure ensembles were consistent and well-defined in each solvent environment, suggesting a high quality representation of the peptide structure in each environment.

Regions of flexibility in a peptide are frequently initiated by Pro or Gly residues, which are known to disrupt defined secondary structure. In Pro residues there are no amide protons available for hydrogen bonding. Whilst the amide proton in Gly residues is available for

hydrogen bonding, the small side chain allows for increased conformational freedom relative to other residues [255, 256]. This creates kinks particularly when multiple Gly or Pro residues are in close proximity. This effect has been illustrated in globular proteins where the presence of high numbers of Gly residues have been implicated in assisting in a reversal of the chain direction [257].

Fallaxidin 3.1(NH_2) contains Gly and Pro residues placed periodically three to four residues apart in the centre of the peptide that can contribute to disorder in this region of the peptide. In DPC micelles, the structure of fallaxidin 3.1(NH_2) has clear kinks in its helical structure at the Pro and Gly residues resulting in distortions in the backbone torsion angles. As they occur every three to four residues, there are insufficient residues to form a full turn of the α helix in between these residues. This creates an overall disruption of the defined helical structure. Collagen also has Gly periodically three to four residues apart in its primary structure which contributes to its ability to form a triple helix with self aggregation of three polypeptide chains [258].

4.3.2 Peptide Profile, Structure and Biological Activity

Litoria fallax is unusual in that it does not secrete a peptide that displays broad spectrum antimicrobial activity, nor a peptide which is a potent inhibitor of nNOS. There are numerous small peptides displaying unique primary structures with unknown biological function. It is also surprising that *L. fallax* does not secrete caerulein, as typically, many frogs of *Litoria* genus (with the exceptions of *L. rubella* and *L. electrica*) produce this peptide [12, 39]. *L. fallax* produces a skin peptide profile (with very limited determined biological function) quite distinct from those of related *Litoria* species. This is also a feature of *L. rubella* whose distinct skin peptide profile does not contain any antimicrobial peptides. Instead it contains a large number of tryptophyllins, many of which have unknown biological activity but are thought to act as neuropeptides [46, 259, 260].

It is not unusual for the secreted peptides to be cytotoxic to the frog: as a consequence, after a period of time on the skin these peptides are deactivated by endoproteases [27]. This often involves removing several N-terminal residues and results in the elimination of biological function [31]. Two peptides, fallaxidin 2.1.1 and 2.2.1 were isolated as part of the skin peptide profile of *L. fallax* and identified as degradation products.

4.3.2.1 Fallaxidin 1 Peptides

Fallaxidin 1 peptides are unique when compared with peptides isolated from other Australian amphibians. Several other species have been shown to produce small peptides exhibiting neuropeptide activity [27]. However to date, the small peptide from *L. fallax* are the first to be identified from Australian amphibians with a Tyr residue at the N-terminus. The activity of these peptides is not known. Fallaxidins 1.1 – 1.3 did not display any activity in smooth muscle contraction, enhancement of lymphocytes proliferation or electrical stimulated guinea pig ileum (GPI) *in vitro* testing. These peptides however, did produce a secondary delayed contraction at the micromolar range upon being washed out of the organ baths. This response was not inhibited by atropine (muscarinic receptor (mAchR) antagonist) [261], L-NNA (NOS inhibitor) [262] or brethylium tosylate (inhibits release of noradrenaline from synaptic nerves) [263]. Thus, it is believed that these peptides are acting via receptors that do not cause release of: (i) acetylcholine (Ach) (eg. CCK_2 receptors [264], neurokinin receptors [265] or orexin receptors [266]); (ii) NO (eg. B_1 receptors [267]); or (iii) noradrenaline from cells or nerve junctions (eg. β_3 adrenoceptors [268]). See Section 6.2.5 for more details. The role of this effect in the animal's host defence is unclear.

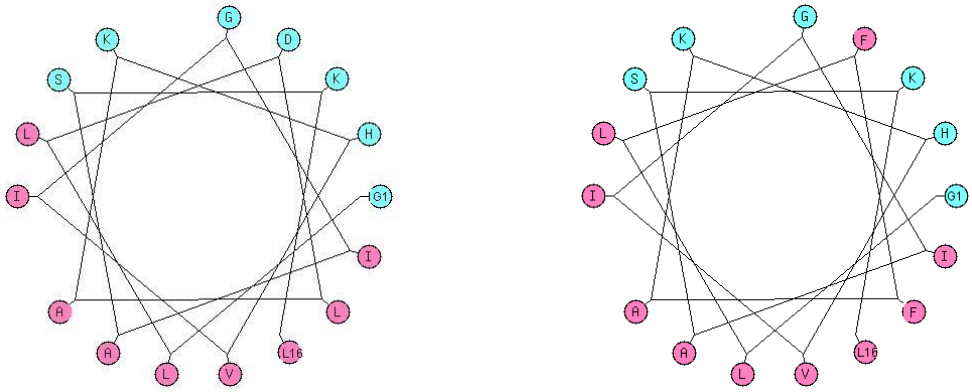
The presence of multiple copies of the fallaxidin 1 peptides encoded in the precursor genes is a characteristic feature of some neuropeptides [269-271]. Thus, it is likely that the fallaxidin 1 peptides have a defined role acting in the central nervous system (CNS). Fallaxidin 1 peptides show some sequence similarities to amphibian opioid peptides; dermorphins [272] and deltorphins [273]. These peptides contain D-amino acids at position 2. This raises the question that the fallaxidin 1 peptides may contain D-amino acids. More investigation is required to determine this.

Dermorphin	Y(D-Ala)FGYPS-NH_2
Deltorphin B	Y(D-Ala)FEVVG-NH_2

4.3.2.2 Fallaxidin 2 Peptides

The amphipathic nature of a peptide existing in a helical conformation is generally considered to be a requirement for antibacterial activity. The Schiffer-Edmundson helical wheel diagram is a graphical representation that allows the degree of amphipathicity of a peptide sequence in a helical conformation to be measured [274]. The helical wheel diagram illustrates an end-on-view of the long axis of the peptide with each residue separated on the wheel by intervals of 100° (corresponding to 3.6 residues per turn as in an ideal helix). As the solution structure of fallaxidin 2.1 and 2.2 has not been determined at present, the degree of amphipathicity can be indicated by the wheel projection (Figure 4.20). These peptides are quite hydrophobic and Figure 4.20 suggests that amphipathicity could result from a α helical structure. This, however, does not explain the limited antimicrobial efficacy of these peptides. Some of the most active smaller antimicrobial agents (~ 16 residues) aurein [28] and citropin [35] peptides contain two Lys residues at positions 7 and 8, both of which are necessary for significant antimicrobial activity. It is possible that the substitution of Lys8 with His8 in the fallaxidin 2 peptides has reduced the antimicrobial efficacy of these peptides.

Fallaxidin 2 peptides are narrow spectrum antimicrobial compounds; fallaxidin 2.1 is active against *Leuconostoc lactis* and *Enterococcus faecalis* at a moderate 100 $\mu g.mL^{-1}$, whilst fallaxidin 2.2 is only active against the former. Considering the close structural similarities of these peptides, it is surprising that they do not have the same activity. Despite the governing factors of activity not being entirely understood, it appears that the charged Asp4 (instead of Phe4) in fallaxidin 2.1 is required for activity against *Enterococcus faecalis* in this case.



(a) **(b)**

Figure 4.20: The Schiffer-Edmundson helical wheel projection of fallaxidins **(a)** 2.1 and **(b)** 2.2. Hydrophobic residues are illustrated in pink and hydrophilic residues are illustrated in light blue.

4.3.2.3 Fallaxidin 3 Peptides

The native fallaxidin 3.1 displays improved antimicrobial activity against Gram-positive organisms when compared to the fallaxidin 2 peptides. The Schiffer-Edmundson helical wheel projection for fallaxidin 3.1 (Figure 4.21) suggests that the peptide is amphipathic and contains a charge of +1. The modified fallaxidin 3.1(NH_2) (+2) displays remarkably improved antimicrobial activity against Gram-positive organisms (Table 4.10). Many previous studies have revealed that the positive charge of a peptide is related to its antimicrobial activity, with peptides that have a higher positive charge displaying greater activity [50, 55, 56]. It is therefore interesting that the animal secretes this peptide with a free acid instead of the more active amide, given that a large number of amphibian peptides are in fact post-translationally modified to the amide.

The fallaxidin 3.1(NH_2) secondary structure indicates that despite the disruption in the helical structure of the peptide due to kinks created by Pro and Gly residues, the peptide displays amphipathicity consistent with the helical projection (Figure 4.21). It is possible that the kinks in the secondary structure impart some flexibility to the peptide, allowing it to interact efficiently with the bacterial membrane leading to lysis of the cell. Partial helical structures and random conformations have been reported for several antimicrobial peptides, in particular dermaseptin S3, which has been shown to exhibit conformational flexibility [275]. These flexible antimicrobial peptides display moderate activity against many organisms but when their overall structure is stabilised to more ordered

conformations, a decrease in the MIC values results [275]. It is possible that the flexibility observed in fallaxidin 3.1 may account for the moderate overall activity against certain Gram-positive bacteria.

Generally, a wide spectrum amphibian antibiotic agent displays MIC values in the low micromolar concentration range. Fallaxidin 3.1 has low MIC micromolar concentrations against some of the Gram-positive organisms tested. Fallaxidin 3.1 is particularly effective at inhibiting the growth of *Leuconostoc lactis*, *Streptococcus uberis* and *Micrococcus luteus*. *L. lactis* is often found on plant surfaces, while *S. uberis* and *M. luteus* mainly grow on mammalian skin [276]. The optimal temperature for growth of each organism [277] corresponds to the subtropical climate of the primary habitat of *L. fallax*. *L. fallax* resides nearby to water and among emergent vegetation, thus the animal is likely to be in contact with sources of the aforementioned organisms from plant material and mammals frequenting the water sources. The presence of fallaxidin 3.1 in the skin secretion provides *L. fallax* with protection against infection from these bacteria.

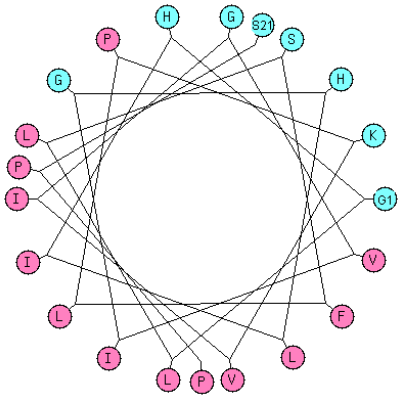


Figure 4.21: The Schiffer-Edmundson helical wheel projection of fallaxidin 3.1. Hydrophobic residues are illustrated in pink and hydrophilic residues are illustrated in light blue.

4.4 Experimental

4.4.1 Collection of Skin Secretions

The specimens of *Litoria fallax* were provided by Stuart Blackburn (Modbury, South Australia, Australia) and the skin secretions collected with assistance from Assoc. Prof. Michael J. Tyler (School of Earth and Environmental Sciences, The University of Adelaide). The animals were held by their hind legs and the skin was washed with deionised water. The dorsal skin glands were stimulated using a bipolar electrode of 21 G platinum connected to a Palmer Student Model electrical stimulator. The electrode was used in a circular motion at 2 V and a pulse rate of 4 ms. The generated secretion was washed off using deionised water and the enzymes denatured with an equal volume of methanol. The mixture was filtered (0.45 μm, Millex HV filter unit) and concentrated to a volume of 1 mL. This procedure was approved by The University of Adelaide Animal Ethics Committee. No animals were harmed during the procedure.

4.4.2 Peptide Secretion Separation by HLPC

The concentrated peptide secretions were separated using HPLC. Separation was achieved using a VYDAC C_{18} HPLC column (5 μm, 300 Å, 4.6 x 250 mm) (Separations Group, Hesperia, CA, USA) that was equilibrated with 10 % acetonitrile/aqueous 0.1 % TFA. The secretions were filtered using a Millex HV filter unit (0.45 μm) and 250 μL of the filtered mixture was injected onto the column. The elution profile was generated by a linear gradient created by a ICI DP 800 data station (ICI Instruments Australia, Melbourne, Victoria, Australia) that controlled two Waters (501 and 510) HPLC pumps (Millipore-Waters Chromatographie, Eschborn, Federal Republic of Germany). The gradient increased the acetonitrile concentration from 10 to 75 % over a period of 30 min at a flow rate of 1 $mL.min^{-1}$. The separated secretions were monitored by UV absorbance at 214 nm using an ICI LC-1200 variable wavelength detector (ICI Instruments Australia, Melbourne, Victoria, Australia).

4.4.3 Mass Spectrometry

The peptides were analysed by ESI-MS using a Micromass Q-TOF 2 orthogonal acceleration TOF mass spectrometer (with mass range to 10,000 Da). The Q-TOF 2 was fitted with an electrospray source in an orthogonal configuration with a Z-spray interface. The peptide samples that were separated by HPLC were diluted in acetonitrile/water (1:1 v/v) and infused into the electrospray source at a flow rate of 5 $\mu L.min^{-1}$. The following conditions were utilised: capillary voltage 3.10 kV; source temperature 80 °C; desolvation temperature 150 °C; and cone voltage 50 V. Optimal fragmentations for MS/MS data were obtained by applying an Ar collision gas energy of between 40 – 80 eV. Sequencing of the peptides from MS/MS spectra was achieved by examination of the **b** and **y** fragmentations and α and β fragmentations in the positive and negative ion modes respectively.

4.4.4 Automated Edman Sequencing

All automated Edman sequencing was conducted by the Department of Molecular Biosciences, The University of Adelaide. An applied Biosystem 492 Procise Sequencer equipped with a 900A data analysis module, using standard procedure [278] was used to analyse the peptides. The peptide, in aqueous acetonitrile (90 %) was absorbed onto a disc immobilon film that had been pretreated with bioprene in ethanol. The disc was pierced several times using a razor blade to help with the solvent flow.

4.4.5 Materials

4.4.5.1 Synthetic Peptides

The peptides were synthesised with L-amino acids using the standard N-α-Foc method, by GenScript Corp (Piscataway, New Jersey, USA) (for full details refer to Maeji *et al.*) [279].

4.4.5.2 NMR Materials

All samples used for NMR structure determination had greater than 90 % purity, determined by HPLC and ESI-MS. Deuterated TFE and DPC was purchased from Cambridge Isotope Laboratories (Andover, MA, USA). Water was filtered using a Milli-Q system (Millipore).

4.4.6 Sample Preparation

The NMR structure determination of amphibian peptides was performed in aqueous d_3-TFE and micelle solutions. Aqueous TFE solutions were prepared by dissolving fallaxidin 3.1 (4.8 mg, 2.19 x 10^{-6} moles) in a TFE/H_2O (1:1 v/v) mixture to produce a final volume of 0.7 mL. The pH was recorded using a Eutech Cyberscan pH 500 Meter with an AEP 3311 glass-body pH probe (183 x 4 mm, thin stem). A single small crystal of sodium 2,2-dimethyl-2-silapentane-5-sulphonate (DSS) was added as a reference.

For the structure determination of fallaxidin 3.1(NH_2) in micelles, the peptide (7.67 mg, 3.5 x 10^{-6} moles) and 40 molar equivalents of DPC (54.56 mg, 1.4 x 10^{-4} moles) were dissolved in an aqueous solution containing 10 % D_2O and 50 mM NaH_2PO_4 buffer to produce a final volume of 0.7 mL. The DPC concentration used was in vast excess of the CMC [232]. A peptide/lipid ratio of 1:40 was chosen to give a final solution with approximately one peptide molecule per micelle [234, 236, 237]. The pH was adjusted to 6.0 (using sodium hydroxide) and a single crystal of DSS was added as a reference.

4.4.7 NMR Spectroscopy

All NMR experiments were performed on a Varian Inova-600 NMR spectrometer using a ^1H frequency of 600 MHz and a ^{13}C frequency of 150 MHz. Experiments were run at 25 °C. ^1H NMR spectra were referenced to either methylene protons of the TFE (3.918 ppm) or to DSS (0.0 ppm). In the ^{13}C-^1H HSQC spectra, the ^{13}C F1 dimensions were referenced to TFE $^{13}CH_2$ (60.975 ppm). DSS (0.0 ppm) was used as an internal standard.

2D NMR experiments TOCSY, DQF-COSY and NOESY were acquired in phase-sensitive mode, using time proportional phase incrementations in t_1 [280]. Typically 16 time-averaged scans were acquired per increment, totalling 256 increments per experiment. The FID in t_2 contained 2048 data points spread over a spectral width of 6000.2 Hz[1]. NOESY spectra were acquired using mixing times of 150 ms. TOCSY pulse sequences included a 70 ms spin-lock. Presaturation was used to suppress water resonances. The ^{13}C-^{1}H HSQC experiments were recorded with an interpulse delay of $1/2J_{CH}$ = 3.6 ms corresponding to J_{CH} = 140 Hz. In the ^{13}C-^{1}H HSQC experiment 256 increments were used, each consisting of 16 scans. The FID in the directly detected ^{1}H, F_2 dimension consisted of 4096 data points and spectral width of 6600.1 Hz. A spectral width in the ^{13}C, F_1 dimension of 21117.0 Hz was also used.

For fallaxidin 3.1(NH_2) in DPC micelles, all NMR experiments included 200 increments, each consisting of 64 time-averaged scans. The FID in t_2 consisted of 2048 data points spread over a spectral width of 6999.7 Hz[1]. For the TOCSY experiment, a spin-lock of 60 ms was used, whilst the NOESY experiment used a mixing time of 100 ms. For the ^{13}C-^{1}H HSQC experiment, the FID in the F_2 and F_1 dimensions had spectral widths of 6999.7 and 21114.5 Hz respectively.

All resulting 2D NMR experiments were processed on a Sun Microsystems Ultra Sparc 1/170 workstation using VNMR software (VNMRJ, version 1.1D). The data matrices were multiplied by a Gaussian function in both dimensions and zero-filled to 2048 data points before Fourier transformation (^{13}C-^{1}H HSQC: 4096 data points in F_1). The processed 2D NMR matrices contained 2048 x 2048 real data points (^{13}C-^{1}H HSQC: 4096 x 4096).

$^{3}J_{NH\alpha H}$ values were measured from high-resolution 1D ^{1}H NMR spectra acquired using 0.125 Hz per point digital resolution. Dihedral angles were restrained to: $^{3}J_{NH\alpha H} < 5$ Hz, $\phi = -60° \pm 30°$ and $5 > {}^{3}J_{NH\alpha H} < 6$ Hz, $\phi = -60° \pm 40°$. For $^{3}J_{NH\alpha H} > 6$ Hz, ϕ angles were not restrained.

4.4.8 Structure Calculations

Cross-peaks in TOSCY and NOESY spectra were assigned using SPARKY software (version 3.106) and a standard sequential assignment method [174]. The NOESY cross-

[1] A 90° pulse was calibrated for each experiment and a 1 s recycle delay was incorporated.

peak volumes were converted to distance restraints using Xu's method [208]. When symmetric pairs of cross-peaks were present, the larger peak volume was used and converted to a distance restraint [205].

Structures of peptides were generated using ARIA (version 1.2) applied with CNS Solve (version 1.1) (Section 3.8). The standard RMD and SA method of ARIA was used [211] with the inclusion of floating stereo-specific assignments to enable the assignment of distance restraints relating to methylene and isopropyl group resonances [281]. Each ARIA run comprised of eight iterations. The structures of the peptides were calculated using the standard default ARIA parameters in the run.cns task file. Better convergence was achieved using the "ambigcutoff" parameters based on those used by Pari *et al*. [282] and by doubling the number of steps in the SA protocol as described by Kang *et al.* [220]. Sixty structures were calculated in the final iteration and the twenty lowest potential energy structures were selected for analysis. The final 3D structures was displayed using VMD software (version 1.8.2) [283] and MOLMOL (version 2k.2) [284].

4.4.9 Biological Activity

4.4.9.1 Smooth Muscle Contraction Assay

For full details of the experimental conditions used, see Chapter 6.

4.4.9.2 Lymphocytes Proliferation Studies

For full details of the experimental conditions used, see Chapter 6.

4.4.9.3 Opioid Activity Studies

For full details of the experimental conditions used, see Chapter 6.

4.4.9.4 Antibacterial Testing

Antibacterial testing was performed by the Microbiology Department of the Institute of Medical and Veterinary Science (Adelaide, South Australia, Australia). The peptides were applied to microorganism coated agarose plates and measurements of the resulting inhibition zones were recorded [285]. All activities were recorded as MIC values ($\mu g.mL^{-1}$). These values are the minimum inhibitory concentrations that were required to inhibit the growth of the microorganisms.

4.4.9.5 cDNA Sequencing

For full details of experimental conditions used, see Chapter 7.

Chapter 5 Peptides That Inhibit Neuronal Nitric Oxide Synthase

5.1 Introduction

5.1.1 Nitric Oxide

Originally discovered as the unidentified “endothelium-derived relaxing factor” by Furchgott and Zawadzki in 1980 [286], it was not until 1987 that it was identified as NO [287-289]. From this time, NO has become a major focus of biological research in physiology, health and disease. NO is a ubiquitous biological signalling molecule that mediates a wide range of activities in both eukaryote and mammalian physiology. In eukaryotes, NO is involved in controlling a diverse range of processes including swimming motion in tadpoles [290], gill muscle function in clams [291] and the generation of light in fireflies [292]. NO also plays a role during courtship in male newts [293] and frog amplexus in *Rana esculenta* [294].

In mammals, NO regulates a variety of biological processes within the immune, circulatory and central nervous systems [295]. The most characterised examples of these physiological processes are the regulation and vasodilation of smooth muscle tone [296-298], neuronal transmission and responses [299, 300], regulation of exocrine secretion [301] and platelet adherence and aggregation [302]. In addition, at low concentrations NO is important as a biological second messenger [303], however, at high concentrations it acts as a cytotoxic agent, able to modulate the immune host defence, particularly against cancerous cells and invading microorganisms [300].

NO is an inorganic free radical, produced in cells with diverse functions [304]. As it is uncharged and hydrophobic [303], NO is able to rapidly diffuse through biological membranes and reach the intracellular environments of nearby cells to exert its wide range of actions [305]. NO differs to conventional neurotransmitters (such as Ach and hormones) as it is released by free diffusion as opposed to exocytosis [297] and its actions involve covalent interactions with targets in contrast to the more common non-covalent ligand-receptor interactions. As a result, NO has an extremely large number of potential biological

targets with the most important biological examples involving interactions with oxygen and superoxides, transition metal ions [306], thiol (R-SH) containing molecules [305], amines [307] and other radicals [303]. It is not surprising with such a diverse range of potential targets, that NO is chemically unstable with a very short half life (3 – 5 s) in biological tissues [305]. Thus, it is not stored and must be synthesised at the site of action. As a consequence, regulation of NO release and action is via control of its synthesis.

Excess production of NO has been shown to be cytotoxic to the body, causing damage to DNA, which can result in energy depletion and cell death [308]. It has also been shown to lead to neurodegeneration, commonly characterised by disease such as Alzheimer's, multiple sclerosis [303] and Parkinson's [300]. In addition, NO has been associated with numerous other disease states including septic shock, stroke, diabetes [306], arthritis, hypertension [309], asthma [308] and migraine headaches [298]. Selective inhibition and regulation of NO synthesis provides a beneficial therapeutic target for the prevention and treatment of many of these diseases.

5.1.2 Nitric Oxide Synthesis

In bacterial cells, NO can be synthesised by either the reduction of nitrite or the oxidation of ammonia [306], however in mammals, synthesis of NO is achieved by a highly regulated enzyme, nitric oxide synthase (NOS). NOS is present in mammals in three distinct isoforms, namely neuronal NOS (nNOS), endothelial NOS (eNOS) and inducible NOS (iNOS) [308]. The nNOS isoform is widely distributed in the cytoplasm of cells in the peripheral neuronal and central nervous systems [304], whilst the eNOS isoform is found membrane bound in endothelial cells. iNOS is predominately distributed in activated macrophages of the immune system [308].

Every functional NOS isoform exists as a protein dimer in the cell [299]. Each nNOS protein subunit is approximately 160 kDa, whereas eNOS and iNOS protein subunits are smaller at approximately 130 kDa [297]. Each isoform of NOS is produced by distinct genes that exist on separate chromosomes. In humans, nNOS is mapped on chromosome 12 [310], eNOS on chromosome 7 [311] and iNOS on chromosome 17 [312]. These genes display an average of 50 % homology between the different isoforms and 90 % homology for the same isoform in different mammalian species [306, 308, 313]. This indicates that

the NOS protein sequences are highly conserved and have a common cellular function. NOS proteins contain a N-terminal catalytic oxygenase domain and a C-terminal reductase domain separated by a 30 amino acid recognition sequence for Ca^{2+} CaM (Figure 5.1) [314]. The N-terminal domain contains binding sites for the iron protoporphyrin IX prosthetic group (heme), tetrahydrobiopterin (BH_4) and L-arginine (L-Arg). Binding sites in series for nicotinamide adenine dinucleotide phosphate (NADPH), flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) are present in the C-terminal domain [297, 306, 308, 314]. NOS has significant sequence homology with cytochrome P_{450} reductase (60 % homology) [297, 308], thus it is not surprising that the same cofactors and cosubstrate binding sites are present.

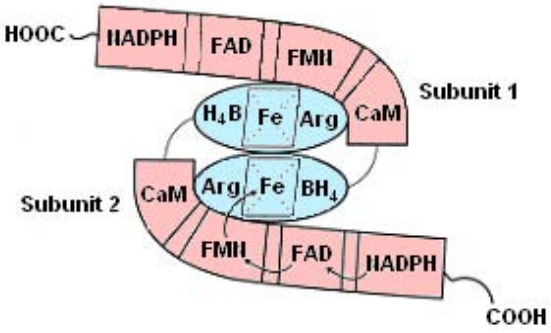


Figure 5.1: Schematic representation of NOS. Electron transfer occurs from the reductase NADPH to the heme group located on the oxygenase domain of the adjacent subunit as shown by the arrows. Oxygenase domains are illustrated in blue and the reductase domain in pink. Adapted from [314, 315].

NOS converts L-Arg to NO and L-citrulline in a two step process involving the five electron oxidation of one of the L-Arg equivalent guanidine nitrogen atoms [297, 306]. The synthesis of NO occurs via the formation of a short lived intermediate N^{ω}-hydroxy-L-arginine (Figure 5.2) [316]. The electrons required to drive the reaction are derived from NADPH bound at the C-terminal reductase domain. NADPH reduces FAD that is bound in close proximity, which subsequently reduces FMN [308, 317]. The binding of Ca^{2+} CaM between the N- and C-terminal domains allows electron movement from the flavin molecules to the heme group [295]. This electron transfer is possible as binding of CaM reorientates the two domains to be more favourably positioned for electron movement [306]. Electron transfer will only occur between the reductase domain and the oxygenase domain of the adjacent subunit (Figure 5.1) [314]. It has been suggested that BH_4 has dual activity in NO synthesis. It may act as an allosteric effector, promoting NOS dimerisation

and stabilising the structure, and in addition, it may be involved in protecting the enzyme from NO feedback inhibition [318, 319]. The cofactor requirements of NOS are important for catalytic activity, for the controlled synthesis of NO and provide possible targets for therapeutic application.

Figure 5.2: The production of NO by NOS. Adapted from [297].

The production of NO by NOS is regulated by a number of cellular and pathophysiological factors. Although the three isoforms share common fundamental structural elements, their activities are regulated differently. nNOS and eNOS are constitutively expressed and their activities are largely controlled by intracellular Ca^{2+} levels [299, 300, 320]. These isoforms are inactive at normal cellular basal Ca^{2+} levels, however an increase in intracellular Ca^{2+} concentration promotes CaM binding and stimulates catalytic activity [313, 320]. This ultimately allows the synthesis of NO to be controlled by physiological stimuli that increase Ca^{2+} levels such as sheer stress [319], neuronal depolarisation and activation of second messengers (such as cGMP) [297, 321]. In contrast, iNOS binds CaM tightly at all times, independently of intracellular Ca^{2+} levels [299, 317]. Regulation of this isoform is controlled by transcription and once induced, iNOS can be active for extended periods of time [306, 309]. iNOS transcription is stimulated in response to pathophysiological situations such as inflammatory responses, infection and disease [303, 308, 319]. Despite this, under normal physiological conditions, iNOS protein levels are either low or undetectable [315].

5.1.3 Calmodulin

CaM is a ubiquitous regulatory protein that activates a wide variety of structurally and functionally distinct enzymes in a Ca^{2+}-dependent manner [322]. CaM binds and regulates enzymes in a variety of physiological processes in response to Ca^{2+} signals, including muscle contraction, cell division, cyclic nucleotide metabolism and glycogen metabolism [323-325]. Over 300 of such proteins have been identified [326]. CaM is highly conserved [327] and present in all eukaryotic cells [328, 329]. It is a small (M_r 16.7 kDa), acidic protein [330], consisting of 148 residues in a single polypeptide chain [327]. The crystal structure of CaM reveals a dumbbell shaped molecule with an overall length of 65 Å. It consists of two globular domains, each with two Ca^{2+} ion binding sites of the helix-loop-helix (EF hand) type motif. These domains are connected by a rigid "central helix", approximately eight turns in length [328, 331]. However, the solution structure has unambiguously demonstrated (^{15}N relaxation experiments [332] and small angle X-ray scattering studies [329]) that the "central helix" is disrupted near its midpoint, adopting an unstructured random coil conformation (Figure 5.3a). As a consequence, it serves as a flexible tether between the two domains. Upon binding Ca^{2+} in an ordered, cooperative manner [333, 334], CaM undergoes a conformational change that results in the exposure of large hydrophobic clefts [328] and enables Ca^{2+} CaM to recognise and bind target proteins with high affinity (Figure 5.3b) [335].

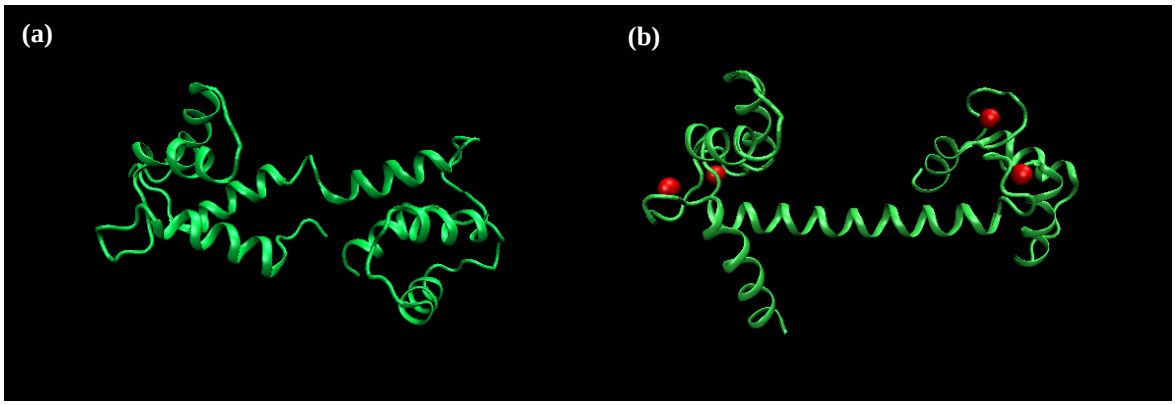


Figure 5.3: 3D representation of **(a)** apoCaM. Adapted from [336], PDB Code: 1cfd. **(b)** Ca_4CaM. Adapted from [331], PDB Code: 1cll. Ca^{2+} ions are illustrated as red spheres.

CaM regulates a variety of proteins involved in the generation of second messengers, such as NOS [337, 338]. As previously discussed, CaM plays a critical role in activating NOS,

by triggering electron transfer from the reductase domain to the heme group [339]. CaM binds to the amphipathic α helical region, which is rich in basic and hydrophobic residues, that connects the two NOS domains [340, 341]. When CaM binds to NOS, it changes the conformation of the reductase domain [342, 343] and increases the rate of electron transfer from NAPDH to the flavins [344]. CaM-dependent regulation provides a means for modulation of NOS and the synthesis of NO.

CaM binds the majority of its physiologically relevant targets when it is fully saturated with Ca^{2+} (Ca_4CaM). Among the most physiological important targets are proteins, however small peptides, hormones, toxins, venoms and drug-like molecules have also been shown to bind [345]. Despite displaying no corresponding sequence homology [346, 347], CaM binds to these targets with high affinity [348]. Generally, CaM binding domains are characterised as short (around 18 residues [349], ~ 15 Å in length [350]), basic, positively charged regions which have the propensity to form amphiphatic α helices [335, 345, 351]. Typically, Ca^{2+} CaM interactions with targets involve strong hydrophobic interactions, with targets containing large hydrophobic residues ("hydrophobic anchors") in conserved positions at 1-5-10, 1-8-10 or 1-14 [345] on one face of the presumed helix. In addition, electrostatic interactions between basic residues in the target peptide and acidic residues in CaM help to stabilise the complex [335].

It has been demonstrated that upon binding targets, Ca^{2+} CaM undergoes significant conformational changes [352]. Two main binding models have been identified; ellipsoidal, collapsed type conformers and extended type conformers, with the former more prevalent [345]. Examples of the collapsed type complexes with Ca^{2+} CaM include myosin light chain kinases (MLCK) [349, 353], CaM-dependent kinase IIα [347] and CaM-dependent kinase-kinase [354]. In these structures, Ca^{2+} CaM adopts a more compact globular conformer, with the N- and C-terminal lobes coming together to form a hydrophobic channel that engulfs the peptide (Figure 5.4a) [345, 355, 356]. The peptide is anchored in the channel at both ends by its large hydrophobic residues [357].

Despite its prevalence, this globular binding mode does not apply to all CaM complexes identified. Some complexes do not collapse into an ellipsoid shape upon peptide binding, instead they remain as extended conformers [345, 358]. The CaM binding domain of the plasma membrane Ca^{2+} pump only binds to the C-terminal lobe of CaM, thus the complex

remains extended (Figure 5.4b) [352]. Variations of this binding model have been identified, with larger binding targets displaying multimeric complexes composed of individual subunits [359, 360]. An example of this is the CaM complex with the Ca^{2+}-activated K^+ channel. A dimeric complex is formed with two K^+ channel fragments bound to two CaM molecules [360].

NOTE:
This figure is included on page 111 of the print copy of the thesis held in the University of Adelaide Library.

Figure 5.4: Representation of globular complex formed between Ca_4CaM (green) and **(a)** the 26 residue peptide fragment of skeletal muscle MLCK (light blue) with residues 3-21 encompassed within Ca_4CaM. Adapted from [353], PDB Code: 2bbm; and **(b)** the binding domain of plasma membrane Ca^{2+} channel (light blue) bound only to the C-terminal lobe of Ca^{2+} CaM. Adapted from [352], PDB Code: 1cff.

5.1.4 Amphibian Peptides that Inhibit nNOS

The majority of Australian amphibians from the genera *Litoria* and *Crinia* studied to date produce at least one active peptide that inhibits nNOS formation of NO in its glandular secretion [14, 27, 361]. Thus far, more than 50 of these peptides have been identified [14, 27]. The nNOS inhibitory peptides are usually active within the micromolar concentration range (Table 5.1). Amphibians have two possible uses for these peptides. Firstly, they may play a role in regulating the animal's physiology as NO is involved in reproduction, sight and gastric modulation in anurans [294, 301, 362]. Secondly, these peptides may form part of the animal's primary defence by interfering with the NO messenger capabilities of its predators. In addition, bacteria has also been shown to produce nNOS [363-365], thus these peptides may also form a defence against invading pathogens.

The active nNOS inhibitory peptides isolated from amphibians fall into three well-defined groups; the aurein/citropin group, the frenatin/dahlein type peptides and the caerin 1 peptides. The first group, the aurein/citropin peptides have multifaceted activity. In addition to their nNOS inhibitory activity, these peptides also display major antimicrobial, anticancer and antifungal activity [35, 366]. These peptides carry positive charges and are short, linear, amphipathic α helices. Among the most active is aurein 2.3 (*L. aurea*, 1.8 μM). Group 2, the frenatin/dahlein type peptides are characterised by a C-terminal free acid and a Lys-X-Lys motif (or Lys-X-Lys-Y-Lys motif, where X and Y are Leu, Pro or Ser) near this terminus [14]. These peptides show no significant anticancer or antimicrobial activity [14, 27]. The nNOS inhibitory activity increases with an increased positive charge and the dahlein 5 peptides are the most active (dahlein 5.6, *L. dahlii*, 1.6 μM). It has been suggested that these peptides are α helical at the N-terminus and unstructured at the C-terminus [14]. The final main group are those caerin 1 peptides which possess an important hydrophobic Phe at position 3. These peptides adopt a bent helical structure [12, 67]. Like group 1, group 3 peptides have multifaceted activity and are among some of the most cytotoxic of all the *Litoria* peptides, displaying widespread antimicrobial, anticancer, fungicidal and antiviral (including HIV) activity [14, 27]. Among the most active of this group is caerin 1.8 (*L. chloris*, 1.7 μM). In addition, there is a group of miscellaneous peptides, which includes peptides that cannot be characterised by the three groups.

Table 5.1: nNOS inhibition of selected amphibian peptides in order of activity within the inhibitory group [14, 27, 41, 361, 367-371].

Name	Species	Sequence	IC_{50} (μM)	Hill Slope
Inhibitory Group 1				
Citropin 1.1.2	*L. citropa*	DVIKKVASVIGGL-NH_2	42.4	1.5
Aurein 1.1	*L. aurea*	GLFDIIKKIAESI-NH_2	33.9	2.0
Citropin 1.1.3	*L. citropa*	GLFDIIKKVAS-NH_2	22.6	2.2
Lesueurin	*L. lesieuri*	GLLDILKKVGKVA-NH_2	13.2	1.5
Signiferin 2.1	*C. signifera*	IIGHLIKTALGMLGL-NH_2	16.6	2.4
Uperin 3.6	*U. mjobergii*	GVIDAAKKVVNVLKNLF-NH_2	4.4	1.5
Citropin 1.1	*L. citropa*	GLFDVIKKVASVIGGL-NH_2	8.2	1.6
Aurein 2.2	*L. aurea*	GLFDIVKKVVGALGSL-NH_2	4.3	2.5
Aurein 2.4	*L. aurea*	GLFDIVKKVVGTLAGL-NH_2	2.1	3.1
Aurein 2.3	*L. aurea*	GLFDIVKKVVGIAGSL-NH_2	1.8	1.7
Inhibitory Group 2				
Signiferin 3.1	*C. signifera*	GIAEFLNYIKSKA-NH_2	81.2	3.4
Citropin 2.1	*L. citropa*	GPIGSIGKALGGLLVDVLKPKL-OH	31.2	1.5
Signiferin 4.3	*C. signifera*	GFADLFGKAVDFIKSRV-NH_2	16.6	2.4
Splendipherin	*L. splendida*	GLVSSIGKALGGLLADVVKSKGQPA-OH	9.1	1.3
Aurein 5.2	*L. aurea*	GLMSSIGKALGGLIVDVLKPKTPAS-OH	7.7	1.4
Frenatin 3	*L. infrafrenata*	GLMSVLGHAVGNVLGGLFKPKS-OH	6.8	1.4
Caerin 2.6	Hybrid[1]	GLVSSIGKVLGGLLADVVKSKGQPA-OH	6.6	1.9
Dahlein 5.1	*L. dahlii*	GLLGSIGNAIGAFIANKLKP-OH	3.2	2.2
Dahlein 5.2	*L. dahlii*	GLLGSIGNAIGAFIANKLKPK-OH	2.6	4.6
Dahlein 5.6	*L. dahlii*	GLLASLGKVFGGYLAEKLKPK-OH	1.6	2.1
Dahlein 5.3	*L. dahlii*	GLLASLGKVLGGYLAEKLKP-OH	1.4	3.0
Inhibitory Group 3				
Caerin 1.10	*L. chloris*	GLLSVLGSVAKHVLPHVVPVIAEKL-NH_2	41.0	0.6
Caerin 1.1	*L. splendida* *L. caerulea* *L. gilleni*	GLLGVLGSIAKHVLPHVVPVIAEHL-NH_2	36.6	1.4
Caerin 1.20	Hybrid[1]	GLFGILGSVAKHVLPHVIPVVAEHL-NH_2	27.2	1.5
Caerin 1.6	*L. chloris*	GLFSVLGAVAKHVLPHVVPVIAEKL-NH_2	8.5	1.7
Caerin 1.9	*L. chloris*	GLFGVLGSIAKHLLPHVVPLIAEKL-NH_2	6.2	2.2
Caerin 1.19	*L. gracilenta*	GLLSVLGSVAKHVLPHVVPIIAEKL-NH_2	4.1	4.7
Caerin 1.19.3	*L. gracilenta*	GSVAKHLLPHVAPIIAEKL-NH_2	Inactive	
Caerin 1.8	*L. chloris*	GLFKVLGSVAKHLLPHVVPVIAEKL-NH_2	1.7	3.6
Caerin 1.8.1	*L. chloris*	GSVAKHLLPHVVPVIAEKL-NH_2	Inactive	
Inhibitory Group 4				
Fallaxin 3.1	*L. fallax*	GLLSFLPKVIGVIGHLIHPPS-NH_2	15.4	3.9
Dahlein 4.2	*L. dahlii*	GLWQFIKDKIKDAARTGLVTGIQS-NH_2	11.1	2.4
Brevinin 1R	*R. ridibunda*	FFPAIFRLVAKVVPSIICSVIKKC-OH	3.9	5.3
Deserticolin 1	*C. deserticola*	GLADFLNKAVGGKVVDFVKS-NH_2	2.4	1.6

[1] Hybrid frog from female *L. splendida* and male *L. caerulea* [372]

It was thought that these amphibian peptides inhibited nNOS by interacting with its regulatory protein Ca^{2+} CaM. Detailed *in vitro* studies of these peptides showed observed Hill slopes [373] to be greater than one (Table 5.1). A Hill slope equal to one indicates that a single active enzyme molecule is interacting with a single substrate and is expected if the peptide is bound directly to the active L-Arg substrate site on the heme group of nNOS. However, as seen here, Hill slopes greater than one are indicative of positive coordination and suggest that another non-competitive interaction is causing the inhibition of nNOS. This may be due to the peptides binding to CaM and inducing a conformational change in the protein that prevented it associating with nNOS. Experimentally, this was supported as the addition of excess CaM to the peptide-inhibited nNOS resulted in a partial recovery of nNOS activity [41]. Furthermore, these peptides also inhibited the function of CaM-regulated calcineurin [41].

5.1.4.1 Dahlein 5 Peptides

Litoria dahlii is an example of an Australian amphibian that produces several nNOS inhibitory peptides, namely the dahlein 5 group of peptides. These peptides strongly inhibit the formation of NO by inhibiting nNOS at micromolar concentrations (Table 5.1) [41]. *L. dahlii* is dependent almost entirely on chemical defences and the dahlein 5 peptides act as a frontline defence, by inhibiting the formation of NO through nNOS of both itself and its predators [36]. The dahlein 5 peptides belong to the nNOS inhibitory group 2; containing a C-terminal CO_2H group, two Lys residues near this terminus as a Lys-X-Lys motif and positive charge [27]. Investigation of these dahlein 5 peptide complexes with CaM may provide further understanding into the binding properties, the mode of action of nNOS inhibition and the potential to assist in pharmaceutical design of improved NOS inhibitors.

The aims of this study were to: (i) determine the solution structure of a member of the dahlein 5 peptide group, namely dahlein 5.6 and to determine if this peptide has the propensity to form an amphiphatic α helix; (ii) confirm whether dahlein 5.6 interacts with CaM to inhibit nNOS; and (iii) investigate the binding mode of the complexes in order to gain information about the non-covalent interactions between the binding molecules.

5.2 Results

5.2.1 Solution Structure of Dahlein 5.6

5.2.1.1 NMR Assignment

For dahlein 5.6, the NMR spectra were recorded in the membrane mimicking solvent TFE/H_2O (1:1 v/v) and also in DPC micelles. The proton resonances were assigned from the TOCSY, COSY and NOESY spectra using the sequential assignment procedure outlined by Wüthrich [174]. In aqueous TFE there was a reasonable amount of overlap observed in the cross peaks of the TOCSY for the amino acids. Amide resonances for Gly7 and Gly11 as well as Gly12 and Ala15 were coincidental. Differentiation between spin systems was achieved through the use of the COSY spectrum and $d_{\alpha N}$ cross-peaks in the NOESY spectrum. The overlap observed made sequential assignment difficult, however the amide region of the NOESY spectrum provided assistance with sequential assignment of the amide proton chemical shifts. The cross section of the amide region can be seen in the partial TOCSY and NOESY spectra in Figure 5.5. The d_{NN} NOE cross-peaks enabled assignment and differentiation of the dahlein 5.6 residues. In DPC micelles, all of the resonances for dahlein 5.6 were easily assigned with assistance of the NOESY and COSY spectra, with no coincidental peaks (Figure 5.6).

A Pro residue (Pro20) was present in the structure of dahlein 5.6, however no additional cross-peaks were observed in the TOCSY or NOESY spectra for either solvent environment. This indicated that only one conformer of Pro is present in the secondary structure of the peptide. Additional peaks were observed in the amide cross section of the TOCSY spectrum (Figure 5.5) for aqueous TFE. These peaks corresponded to cross-peaks of the ε-NH_3^+ protons of Lys19 and were characterised by very strong cross-peaks corresponding to the ε-CH_3 protons and the absence of a cross-peaks for the αH resonances.

The αC resonances were assigned from the ^{13}C-^{1}H HSQC spectrum, focusing on the αH ^{1}H/αC ^{13}C region of the spectra (Figure 5.7 and Figure 5.8). Table 5.2 presents a summary of all the ^{1}H and ^{13}C chemical shifts for dahlein 5.6.

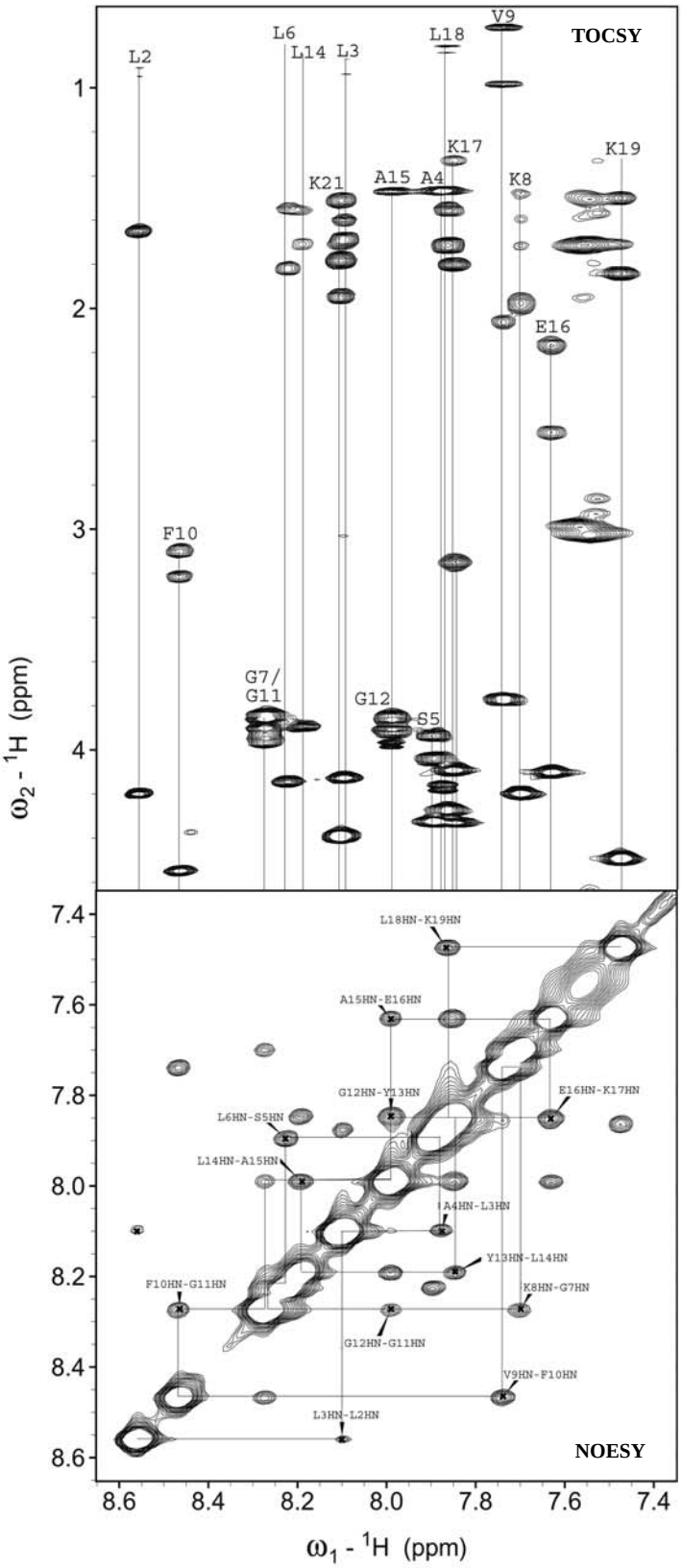


Figure 5.5: Partial TOCSY and NOESY spectra of dahlein 5.6 in TFE/H_2O (1:1 v/v). In the TOCSY spectrum, vertical lines connect resonances in the same spin system. NOEs between sequential amide protons are indicated in the NOESY spectrum.

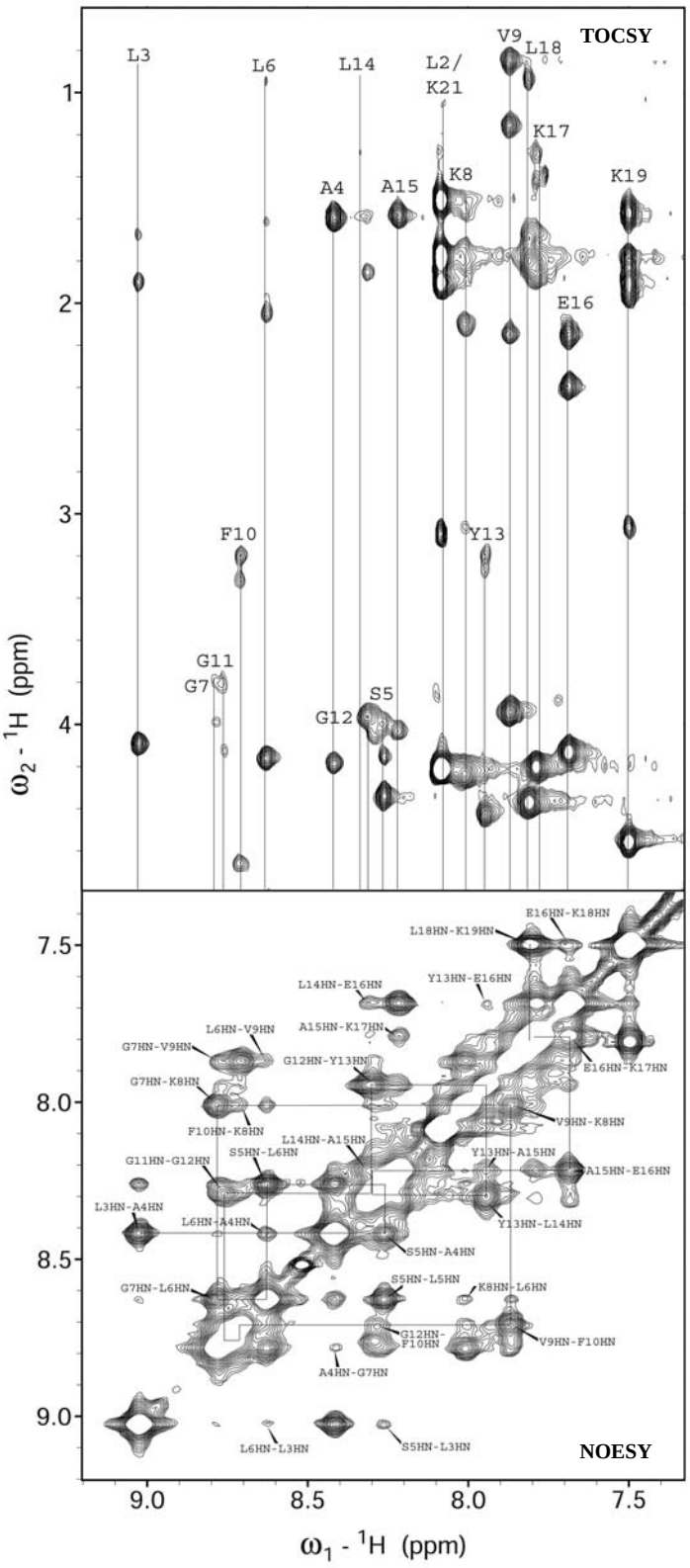


Figure 5.6: Partial TOCSY and NOESY spectra of dahlein 5.6 in DPC micelles. In the TOCSY spectrum, vertical lines connect resonances in the same spin system. NOEs between sequential amide protons are indicated in the NOESY spectrum.

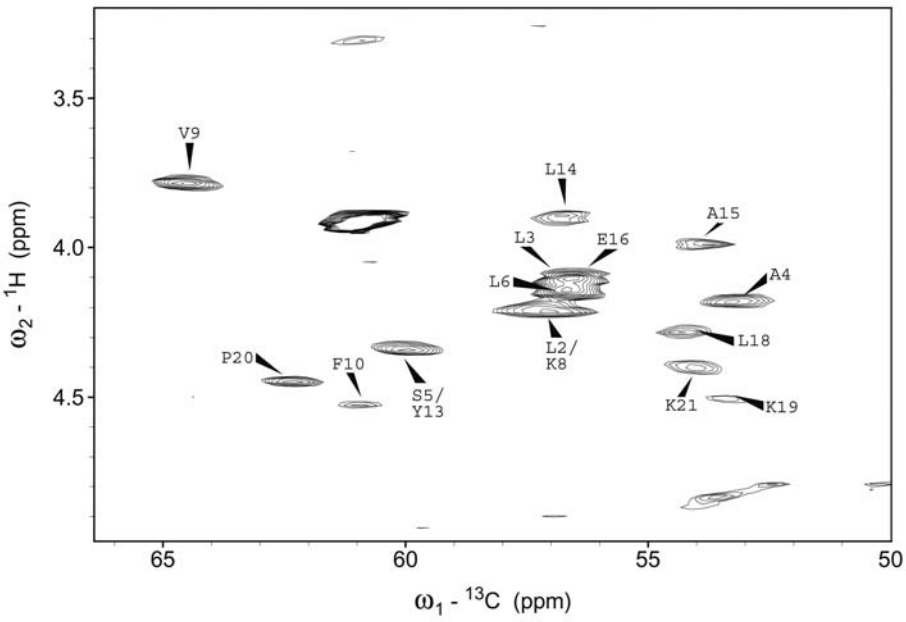


Figure 5.7: Partial ^{13}C-^{1}H HSQC spectrum of dahlein 5.6 in TFE/H_2O (1:1 v/v) showing the αH ^{1}H/αC ^{13}C region. Only the αH-αC connectivities are indicated.

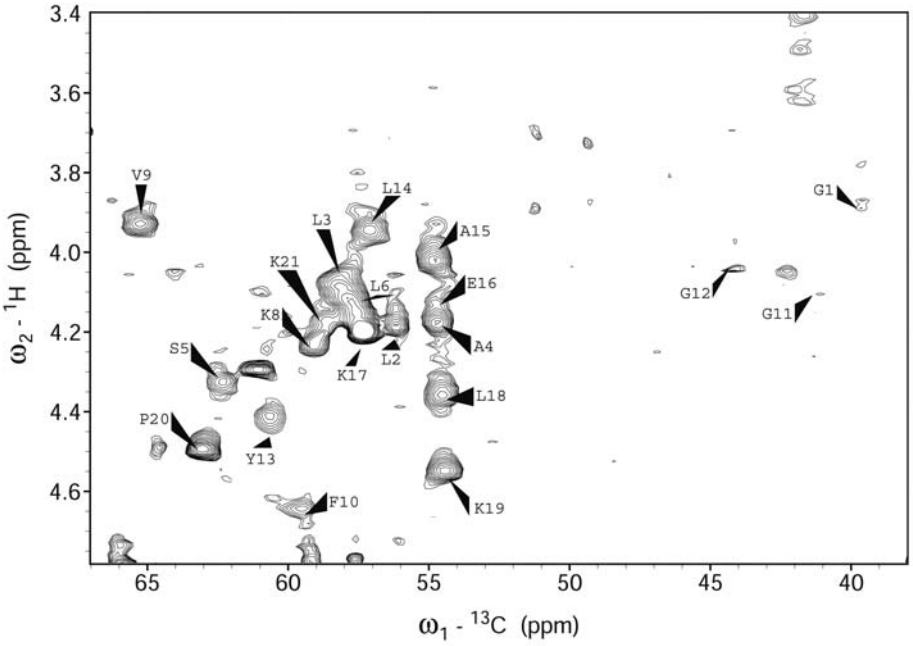


Figure 5.8: Partial ^{13}C-^{1}H HSQC spectrum of dahlein 5.6 in DPC micelles showing the αH ^{1}H/αC ^{13}C region. Only the αH-αC connectivities are indicated.

Table 5.2: ^{1}H and ^{13}C chemical shifts for dahlein 5.6 in TFE/H_2O (1:1 v/v), pH 3.03, 25 ºC. n.o. indicates resonance was not observed.

Residue	**Chemical Shift**				
	HN	**αH**	**βH**	**Other H**	**α ^{13}C**
Gly1	n.o.	3.94, 3.87			n.o.
Leu2	8.56	4.20	1.65	γ-CH n.o	57.14
				δ-CH_3 0.94, 0.91	
Leu3	8.10	4.13	1.69	γ-CH 1.60	56.52
				δ-CH_3 0.94, 0.87	
Ala4	7.88	4.17	1.49		53.25
Ser5	7.90	4.33	4.04, 3.94		59.95
Leu6	8.23	4.15	1.82, 1.75	γ-CH 1.55	56.66
				δ-CH_3 0.88, 0.84	
Gly7	8.27	3.90, 3.86			n.o.
Lys8	7.70	4.20	1.98	γ-CH_2 1.60, 1.48	57.14
				δ-CH_2 1.72	
				ε-CH_2 2.99	
				ε-NH_3^+ 7.57	
Val9	7.74	3.77	2.07	γ-CH_3 0.99, 0.73	64.56
Phe10	8.47	4.55	3.21, 3.10	H 2,6 7.21	60.92
				H 3,5 7.25	
				H 4 7.20	
Gly11	8.27	3.95, 3.84			n.o.
Gly12	7.99	3.91, 3.86			n.o.
Tyr13	7.85	4.33	3.15	H 2,6 6.78	59.97
				H 3,5 7.08	
Leu14	8.19	3.90	1.71	γ-CH_2 1.56	56.76
				δ-CH_3 0.86	
Ala15	7.99	3.97	1.47		53.89
Qlu16	7.63	4.10	2.17	γ-CH_2 2.56	56.48
Lys17	7.85	4.09	1.80	γ-CH_2 1.33	56.48
				δ-CH_2 1.57	
				ε-CH_2 2.30, 2.86	
				ε-NH_3^+ 7.52	

Table 5.2: continues.

Residue	Chemical Shift				
	HN	**αH**	**βH**	**Other H**	**α^{13}C**
Leu18	7.87	4.27	1.72	γ-CH 1.55	54.27
				δ-CH_3 0.83	
Lys19	7.47	4.49	1.84	γ-CH_2 1.50	53.45
				δ-CH_2 1.71	
				ε-CH_2 3.02	
				ε-NH_3^+ 7.54	
Pro20	-	4.49	2.30, 1.84	γ-CH_2 2.02	62.32
				δ-CH_2 3.84, 3.62	
Lys21	8.11	4.39	1.94	γ-CH_2 1.51	54.05
				δ-CH_2 1.79	
				ε-CH_2 3.03	
				ε-NH_3^+ 7.56	

Table 5.3: ^{1}H and ^{13}C chemical shifts for dahlein 5.6 in DPC micelles, pH 6.00, 25 ºC. n.o. indicates resonance was not observed.

Residue	**Chemical Shift**				
	HN	**αH**	**βH**	**Other H**	**α ^{13}C**
Gly1	n.o.	3.89			39.62
Leu2	8.08	4.18	1.79, 1.73	γ-CH 1.60	56.19
				δ-CH_3 1.08, 1.03	
Leu3	9.03	4.09	1.90	γ-CH 1.68	58.13
				δ-CH_3 1.06, 0.98	
Ala4	8.42	4.18	1.60		54.70
Ser5	8.26	4.34	4.16, 4.00		62.58
Leu6	8.63	4.16	2.05, 1.95	γ-CH 1.61	57.58
				δ-CH_3 0.95	
Gly7	8.78	3.99, 3.82			n.o.
Lys8	8.01	4.24	2.10	γ-CH_2 1.57	59.07
				δ-CH_2 1.83, 1.73	
				ε-CH_2 3.07	
				ε-NH_3^+ 7.23	
Val9	7.87	3.92	2.15	γ-CH_3 1.16, 0.85	65.22
Phe10	8.71	4.66	3.31, 3.20	H 2,6 7.29	59.54
				H 3,5 7.18	
				H 4 7.03	
Gly11	8.71	4.13, 3.82			41.12
Gly12	8.29	4.03			44.14
Tyr13	7.95	4.43	3.28, 3.20	H 2,6 7.15	60.63
				H 3,5 6.83	
Leu14	8.31	3.96	1.85	γ-CH_2 1.59	57.13
				δ-CH_3 0.95, 0.91	
Ala15	8.22	4.03	1.58		54.80
Qlu16	7.69	4.13	2.15	γ-CH_2 2.39	54.75
Lys17	7.79	4.20	1.86	γ-CH_2 1.42, 1.29	57.37
				δ-CH_2 1.65	
				ε-CH_2 2.98, 2.83	
				ε-NH_3^+ 6.83	

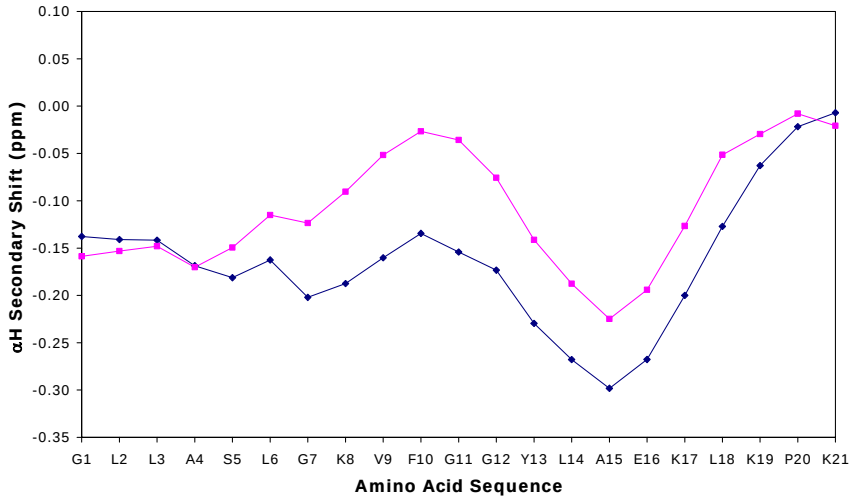
Table 5.3: continues.

Residue	Chemical Shift				
	HN	**αH**	**βH**	**Other H**	**α^{13}C**
Leu18	7.81	4.37	1.81	γ-CH 1.69	54.53
				δ-CH_3 0.94	
Lys19	7.50	4.56	1.95, 1.89	γ-CH_2 1.57	54.44
				δ-CH_2 1.79	
				ε-CH_2 3.06	
				ε-NH_3^+ n.o.	
Pro20	-	4.51	2.39, 2.05	γ-CH_2 2.14	63.06
				δ-CH_2 3.90, 3.73	
Lys21	8.08	4.21	1.95, 1.91	γ-CH_2 1.51	58.78
				δ-CH_2 1.78	
				ε-CH_2 3.10	
				ε-NH_3^+ n.o.	

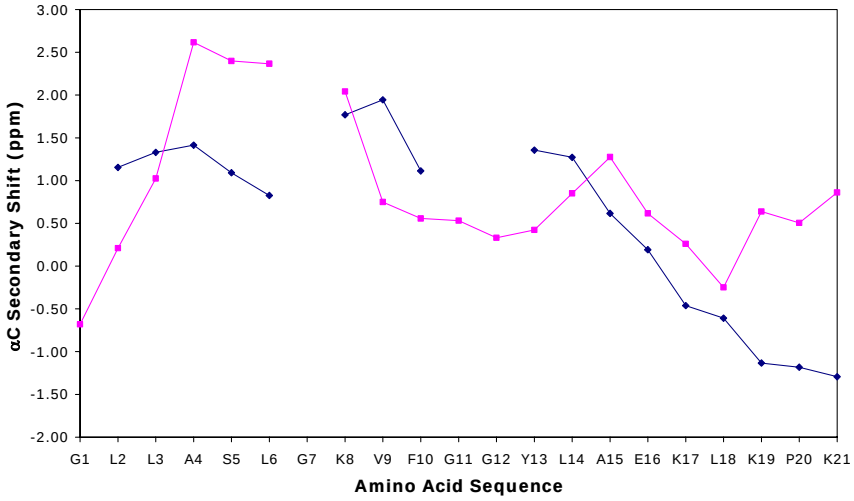
5.2.1.2 Secondary Shifts

The variation of the observed chemical shifts of the amino acid residues from their corresponding random coil values (obtained in water [190]) is referred to as the Δδ. The smoothed (over n = ± 2 residues) αH ^{1}H and αC ^{13}C Δδ and unsmoothed amide ^{1}H Δδ were plotted against the amino acid sequence of dahlein 5.6 (Figure 5.9). As discussed in Section 3.6, upfield shifts in resonances are illustrated by negative Δδ values, whilst positive values indicate a downfield shift in resonances.

(a)



(b)



(c)

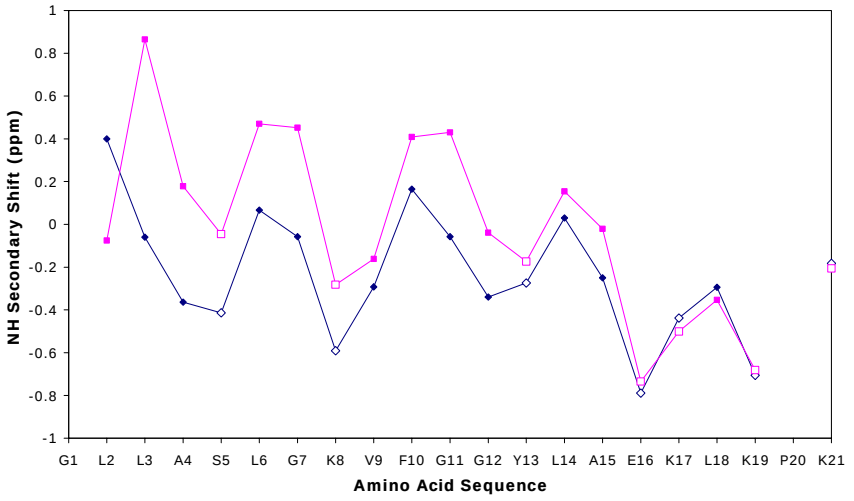


Figure 5.9: (a) αH ^{1}H and **(b)** αC ^{13}C secondary shifts of dahlein 5.6 in TFE/H_2O (1:1 v/v) (blue) and DPC micelles (pink), smoothed over n = ± 2 residues. **(c)** NH ^{1}H secondary shifts of dahlein 5.6, unsmoothed. Negative values indicate an upfield shift from random coil values and positive values indicate a downfield shift. Open symbols indicate hydrophilic residues.

The αH ^{1}H Δδ (Figure 5.9a) have a slight upfield trend from residue 3 to 10 and a strong upfield trend from residue 13 to 18 in both solvents. This is symptomatic of two separate regions of α helical structure either hinged or bent [374]. The near zero regions for the N-terminal residues and two Gly residues (Gly11 and Gly12) suggest that these regions may exist in flexible conformations and implies that there is a possibility of two distinct α helical regions separated by a flexible Gly hinge. This is further supported by the αC ^{13}C Δδ. A general downfield shift is seen, however in aqueous TFE the absence of several central residue chemical shifts means that additional information cannot be gained.

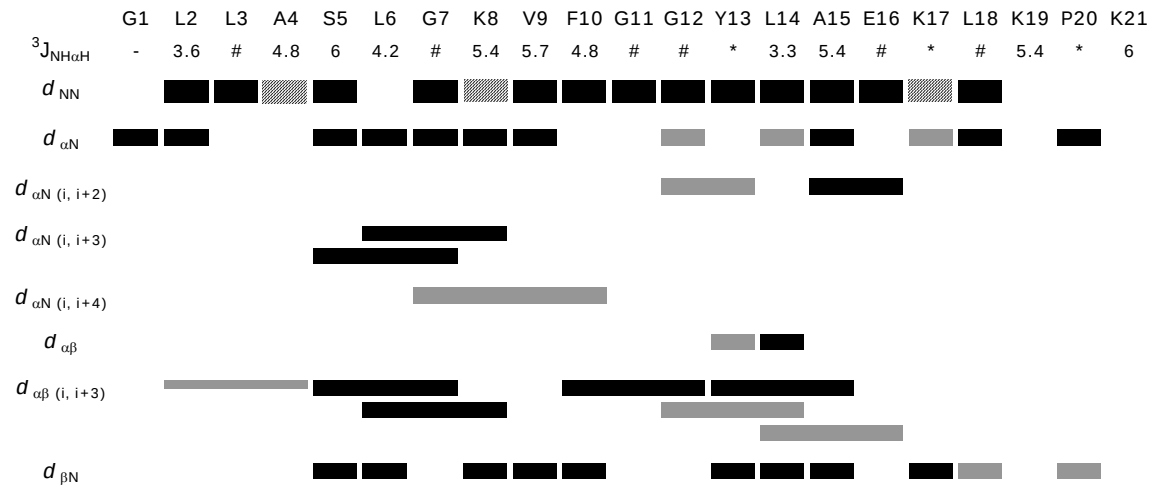
The amide proton Δδ (Figure 5.9c) reveal a three to four periodicity for the entire length of the peptide, where the hydrophilic residues are shifted upfield relative to the hydrophobic residues that are shifted downfield. This characteristic behaviour is observed in amphipathic α helices [197, 198] and is a result of a slight curvature in the helices created by the difference in the hydrogen bond lengths on the hydrophobic and hydrophilic faces of the α helix. In aqueous TFE a general upfield trend at the N-terminus is noted for the amide Δδ values. Once again, this is suggestive of a α helical conformer and is a result of a helical dipole [193]. The variation seen in Figure 5.9c supports the presence of regions of helical secondary structure in the peptide.

5.2.1.3 NOE Connectivities

Figure 5.10 illustrates the diagnostic NOEs obtained from the NOESY spectrum of dahlein 5.6. In aqueous TFE, the absence of medium range NOEs at the N-terminal end of the peptide is consistent with some conformational flexibility in this region, however the $^{3}J_{NH\alpha H}$ coupling constants are consistent with values present in α helical secondary structure. From Figure 5.10, the presence of characteristic α helical NOEs from Ser5 to Ala15 suggests that a region of α helical secondary structure exists in the peptide. In DPC, there are sequential d_{NN}, $d_{\alpha N}$ and $d_{\beta N}$ connectivities present along the majority of the peptide. Additionally, the presence of medium range NOEs three to four residues apart ($d_{\alpha\beta(i,i+3)}$, $d_{\alpha N(i,i+3)}$ and $d_{\alpha N(i,i+4)}$) along the majority of the length of the peptide suggests that in this solvent environment, dahlein 5.6 exists as a α helix.

Due to the presence of Pro20, information about the C-terminal end of the peptide is restricted. The NOE connectivities observed are indications of the extent of the *cis-trans* isomerisation about the Pro imide bond and the preceding residue [248]. As no $d_{\alpha\alpha}$ and $d_{\alpha N}$ connectivities were observed, it can be concluded that the Pro *cis* isomer was not present in this peptide structure (Section 3.5). The presence of strong $d_{\alpha\delta}$ connectivities in the NOESY spectrum indicates that the Pro residue exists in the *trans* conformation.

(a)



(b)

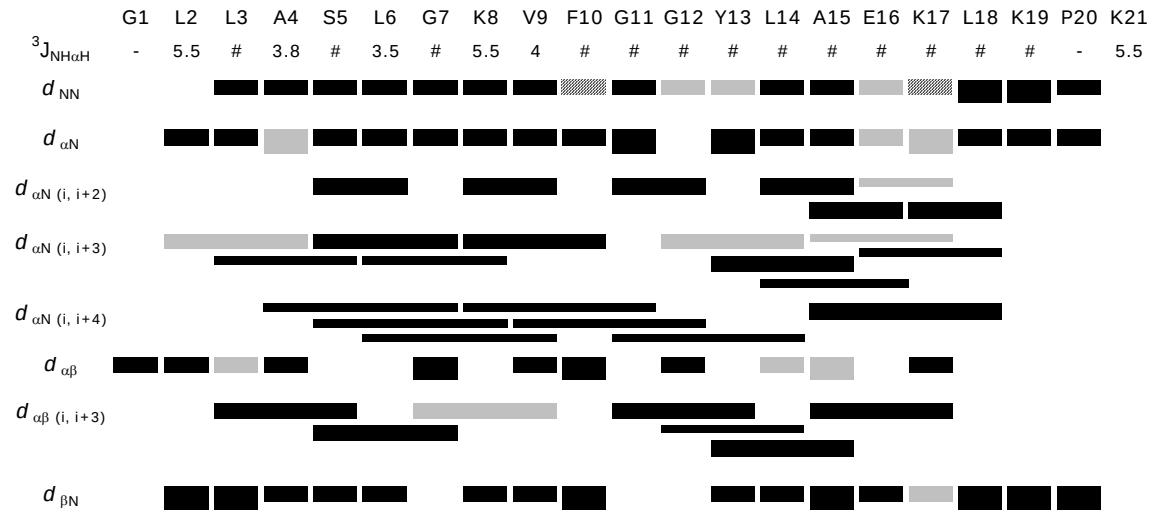


Figure 5.10: A summary of the diagnostic NOEs used in the structure calculations for dahlein 5.6 in **(a)** TFE/H_2O (1:1 by volume) and **(b)** DPC micelles. For Pro where no amide proton is present, NOEs to δH are shown. The thickness of the bar indicates the relative strength of the NOE (strong < 3.1 Å, medium 3.1 - 3.7 Å, weak > 3.7 Å). Grey shaded bars represent ambiguous NOEs. Striped bars represent cross-peaks close to the diagonal. $^3J_{NH\alpha H}$ values are indicated where applicable. # indicates coupling constants that were not resolved, * indicates no coupling constant was detected.

5.2.1.4 Structure Calculations

The NOESY spectra for dahlein 5.6 were fully assigned and the cross-peaks values were quantified and converted to distance restraints as described by Xu *et al.* [208]. A total of 173 and 401 distance restraints were generated from the NOESY spectra for aqueous TFE and DPC samples respectively and are summarised in Table 5.4. In aqueous TFE, very few ambiguous restraints were present, however in DPC, around one fifth of the restraints were ambiguous. Many of these restraints corresponded to side chain protons connectivities.

Table 5.4: The experimental distance restraints obtained from the NOESY spectra of dahlein 5.6 in TFE/H_2O (1:1 v/v) and in DPC micelles.

	No. of Experimental Restraints	
	TFE/H_2O	**DPC**
Sequential NOEs	48	135
Medium-range NOEs	9	100
Long-range NOEs	-	-
Intra-residue NOEs	97	165
Ambiguous NOEs	19	71
Total	173	471

The $^3J_{NH\alpha H}$ coupling constant values were determined from the NH region of the high-resolution 1D ^1H spectrum and can be seen in Figure 5.10. A degree of overlap was observed in this region of the 1D ^1H spectrum, restricting the determination of several coupling constants. Additionally, many peaks were unresolved, consistent with $^3J_{NH\alpha H}$ values less than 5 Hz and α helical in nature. The values for the $^3J_{NH\alpha H}$ coupling constants are consistent with values present in α helical secondary structure. This supports the conclusion from the $\Delta\delta$ plots (Figure 5.9) and the diagnostic NOEs (Figure 5.10) that a region of α helical secondary structure exists in the peptide in both environments.

Sixty structures were initially generated with random backbone torsion angles and these were subjected to RMD and SA calculations. The twenty lowest potential energy structures resulting from these calculations were chosen for analysis. The lowest energy structures for dahlein 5.6 in aqueous TFE and DPC are illustrated in Figure 5.11 and Figure 5.12, superimposed over the backbone atoms of well-defined residues.

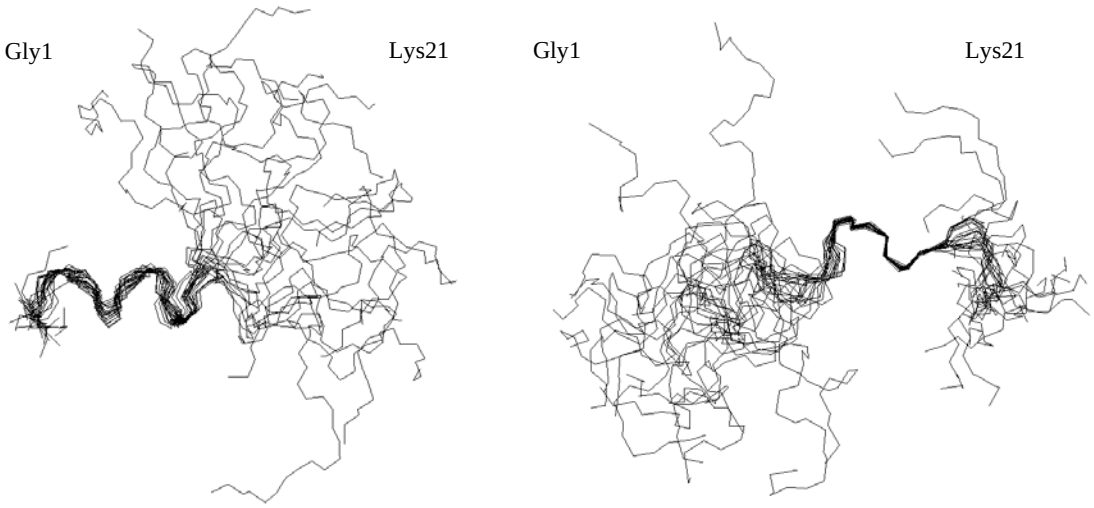


Figure 5.11: The twenty most stable calculated structures of dahlein 5.6 in TFE/H_2O (1:1v/v), superimposed over the backbone atoms of well-defined residues **(a)** 2 – 10 and **(b)** 13 – 16.

Figure 5.11 illustrates that there is some α helical secondary structure in the N-terminal region (Leu2 – Phe10) of the peptide in aqueous TFE environment. This is illustrated more clearly in Figure 5.13. The α helical secondary structure is disrupted by the presence of two Gly residues forming a flexible hinge between the two regions of defined secondary structure (c.f. helix-hinge-helix structure of caerin 1.1 [249]). In DPC however, dahlein 5.6 displays α helical secondary structure along its length with the exception of the terminal residues. There is a bend or flexible hinge evident in the helix about Gly11 and Gly12. The presence of medium range NOE constraints ($d_{\alpha N(i,i+4)}$ and $d_{\alpha\beta(i,i+3)}$) in this region suggests there is defined secondary structure at the two Gly residues. It is most likely that dahlein 5.6 exists as a bent helix (cf. bent helix structure of splendipherin [78]). In aqueous TFE there is no clear secondary structure observed for Lys17 – Lys21. This may be a consequence of the presence of structure disrupting Pro20 [375]. Residues 13 – 16 show some secondary structure, however it appears to be a combination of helical and strand secondary structure (Figure 5.11).

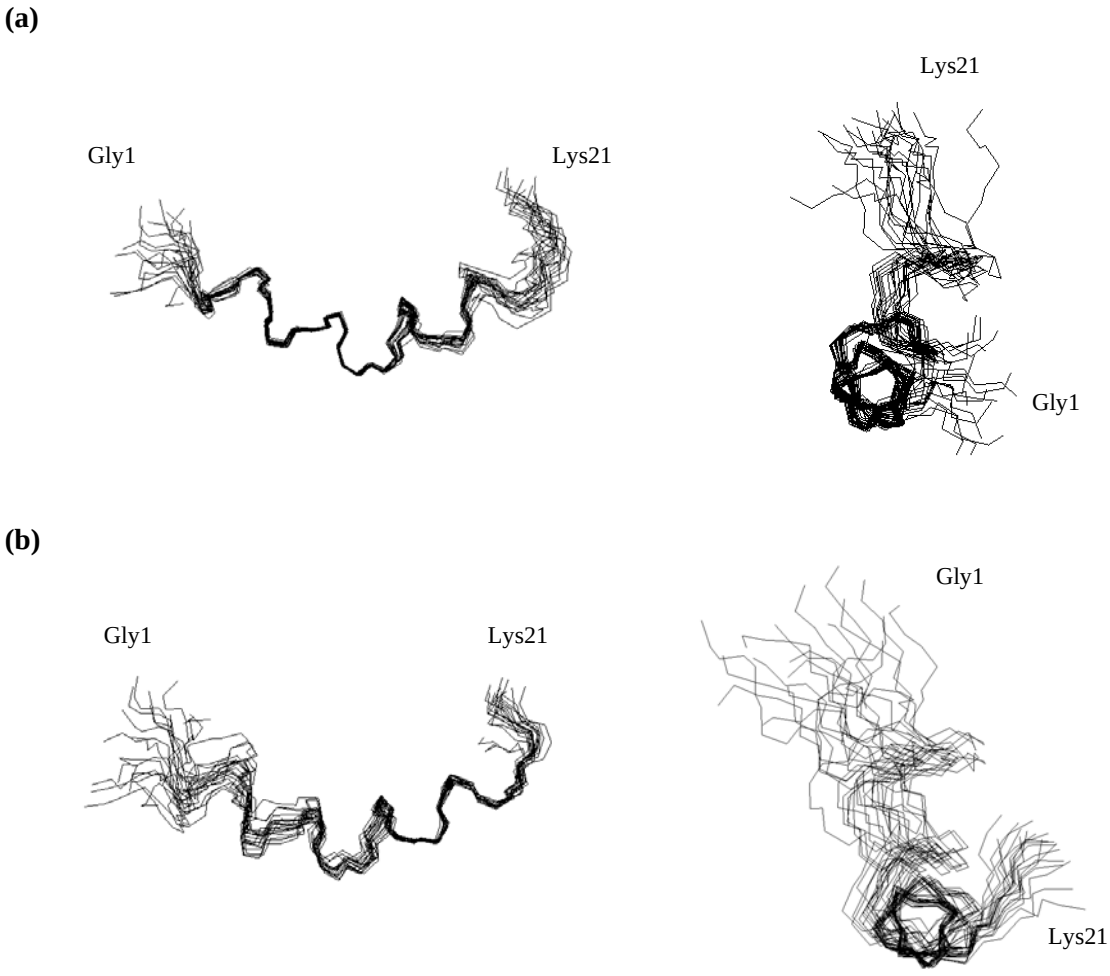


Figure 5.12: The twenty most stable calculated structures of dahlein 5.6 in DPC micelles superimposed over backbone atoms well-defined residues **(a)** 3 – 12 and **(b)** 14 – 19 side on (left) and along the helical axis (right).

The twenty lowest energy structures show high convergence over well-defined residues with RMSD values of backbone atoms in these regions of 0.174 ± 0.060 Å and 0.185 ± 0.065 Å for residues 2 – 10 and 13 – 16 in aqueous TFE respectively. In DPC, geometry RMSD shows comparable values for all well-defined structures. This suggests that for these regions of the peptide, the NMR data has been satisfied. The RMSD values (Table 5.5) for all the atoms of aqueous TFE structures are considerably larger than previously reported structures (c.f. caerin 1.1 in DPC; all backbone atoms – 0.64 ± 0.22 Å and all heavy atoms – 0.90 ± 0.20 Å [249]), possibly highlighting that the peptide exhibits some regions of conformational flexibility in its secondary structure. In DPC, these values are comparable to previously reported structures [78, 249]. The energy and structural statistics are summarised in Table 5.5. The final structures demonstrated only minor

deviations from the idealised covalent geometry (≤ 0.05 Å for bonds and ≤ 5 ° for angles and impropers). Very few distance restraint violations (TFE/H_2O - 3 violations, DPC - 9 violations) greater than 0.3 Å were observed, with maximum violations of 0.41 and 0.51 Å respectively. This implies that the resultant structures satisfy the experimental restraints obtained from the NMR data. Hence, these structures are likely to be a reasonable representation of dahlein 5.6 in membrane mimicking environments.

Table 5.5: The structural statistics for the twenty lowest energy structures of dahlein 5.6 in TFE/H_2O (1:1 v/v) and DPC following RMD/SA calculations.

Energies (kcal mol^{-1})	**TFE/H_2O**	**DPC**
E_{tot}	12.507 ± 0.296	16.062 ± 0.468
E_{bond}	0.124 ± 0.016	0.189 ± 0.025
E_{angle}	8.932 ± 0.071	9.363 ± 0.098
$E_{improper}$	0.188 ± 0.015	0.237 ± 0.017
E_{vdm}	3.254 ± 0.233	6.134 ± 0.406
E_{NOE}	0.008 ± 0.009	0.138 ± 0.106
E_{cdih}	0.001 ± 0.001	0.004 ± 0.008
Well defined residues	2-3, 5-10, 13-16, 18-20	3-12, 14-19

RMSD from mean geometry (Å)				
Heavy atoms of well-defined residues	(5-10)	0.782 ± 0.186	(3-12)	0.595 ± 0.099
Backbone atoms of well-defined residues	(5-10)	0.174 ± 0.060	(3-12)	0.330 ± 0.091
Heavy atoms of well-defined residues	(13-16)	1.179 ± 0.295	(14-19)	0.854 ± 0.024
Backbone atoms of well-defined residues	(13-16)	0.185 ± 0.065	(14-19)	0.263 ± 0.099
All heavy atoms		5.546 ± 1.714		1.182 ± 0.243
All backbone atoms		5.151 ± 1.842		0.909 ± 0.268

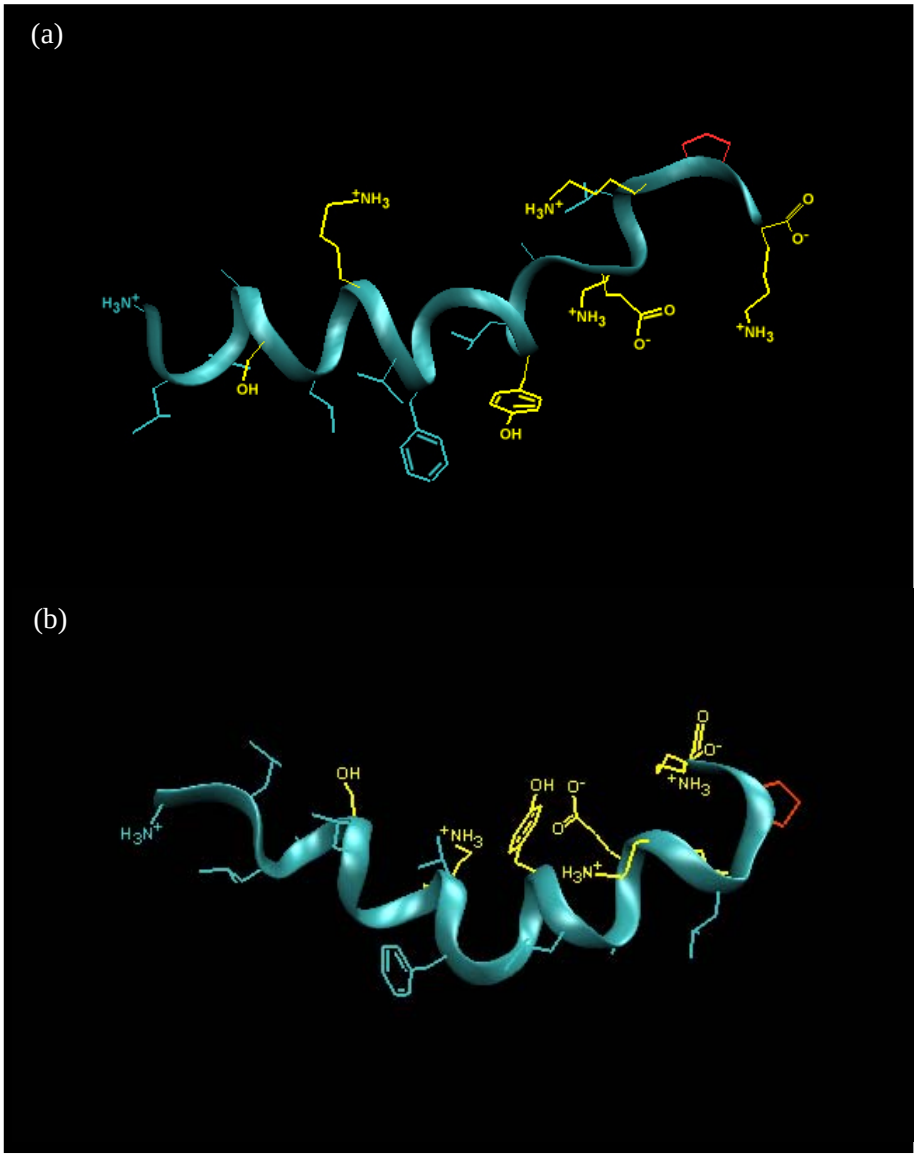


Figure 5.13: The lowest calculated potential energy structure of dahlein 5.6 in **(a)** TFE/H_2O (1:1 v/v) and **(b)** DPC micelles. Yellow indicates residues that are hydrophilic in nature. Pro residues are shown in red.

Analysis of the AOP (S, ϕ and ψ) of the twenty lowest energy structures illustrated that fifteen and sixteen of the twenty-one residues are well-defined (S > 0.9) for aqueous TFE and DPC respectively. Little variation in AOP is expected in regions of well-defined residues. Residues Gly11 and Gly12 do not display well-defined nature in aqueous TFE, which is consistent with a flexible hinge occurring at this position in the peptide. In contrast, these two Gly residues are well-defined in DPC, consistent with the defined secondary structure of a bent helix. The Ramachandran plot (Figure 5.14) for the average ϕ and ψ values illustrate that all the residues fall into the allowed regions. This provides

support for the quality of the calculated structures in representing the solution structures of dahlein 5.6. All of the residues were in favourable or allowable regions.

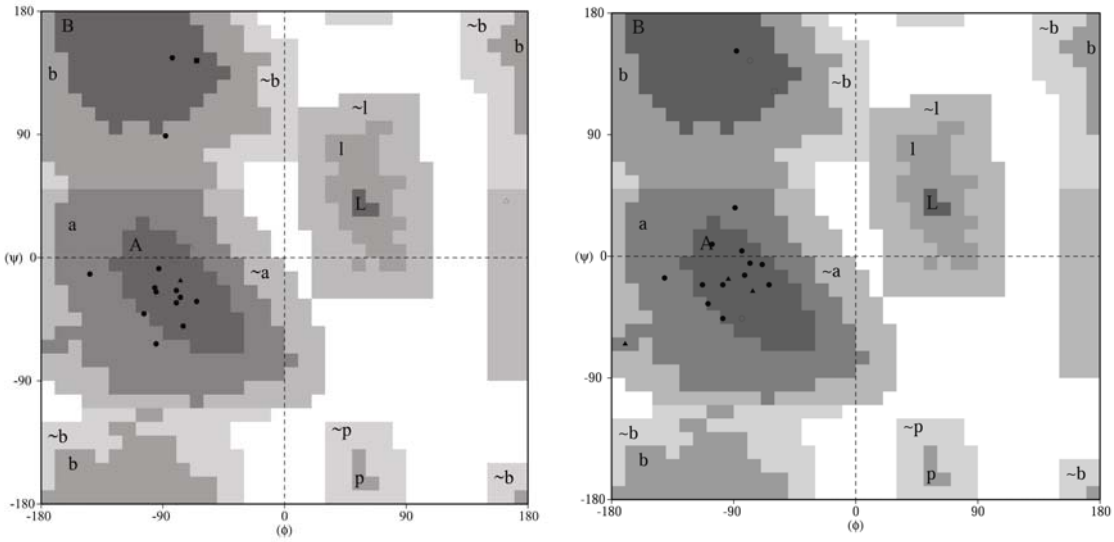


Figure 5.14: Ramachandran plot for the well-defined residues of dahlein 5.6 in **(a)** TFE/H_2O (1:1 v/v) and **(b)** DPC micelles. Favourable regions are labelled A and B for α helical or β strand structures respectively, while allowable and generous regions are labelled a and b or ~a and ~b.

5.2.2 Calmodulin-Dahlein 5.6 Complexes

Dahlein 5.6 was tested for inhibitory actions against nNOS using an assay that measured the [^{3}H]Arg conversion to [^{3}H]citrulline by the nNOS enzyme [41]. Dahlein 5.6 displayed a dose dependent inhibition of nNOS: IC_{50} values and Hill slope of 1.6 ± 0.3 μM and 2.1 respectively. The values obtained are comparable to other amphibian peptides that inhibit nNOS by binding to Ca^{2+} CaM [41, 376].

5.2.2.1 Mass Spectrometry Investigations

It is proposed that dahlein 5.6 inhibits nNOS by interacting with its regulatory protein CaM, causing a conformational change in CaM and preventing its binding to nNOS. It was necessary to ascertain whether dahlein 5.6 binds to CaM to form peptide/protein complexes. CaM is an acidic protein, with an isoelectric point (pI) of approximately 4 [377]. This means that in physiological conditions (pH 7.6), CaM is negatively charged

and as a result, investigations were performed using negative ion ESI-MS. The calculated masses of ions expected in the negative ion ESI-MS of CaM, Ca_4CaM and Ca_4CaM-dahlein 5.6 complex are provided in Table 5.6 as a reference for the possible ions observed and those discussed in this section.

Table 5.6: Calculated masses of the negative ions expected for dahlein 5.6 and CaM ESI-MS investigations.

Charge	***m/z***		
	CaM	**$4Ca^{2+}$ + CaM**	**$4Ca^{2+}$ + CaM + dahlein 5.6**
0	16790	16942	19131
-1	16789	16941	19130
-2	8394	8470	9565
-3	5596	5646	6376
-4	4197	4235	4782
-5	3357	3388	3825
-6	2797	2823	3188
-7	2398	2419	2732
-8	2098	2117	2390
-9	1865	1882	2125
-10	1678	1693	1912
-11	1525	1539	1738
-12	1398	1411	1593
-13	1291	1302	1471
-14	1198	1209	1366
-15	1118	1129	1274
-16	1048	1058	1195

Figure 5.15 shows the negative ion mode ESI-MS spectrum of CaM in ammonium acetate at pH 7.6. Despite extensive dialysis before the analysis, some evidence of binding of Ca^{2+} ions is still apparent once the spectrum has been transformed (illustrated by the additional higher mass peaks). Regardless of the trace presence of Ca^{2+} ions, the predominant species in the spectrum is the Ca^{2+}-free CaM (apoCaM). There is a charge distribution pattern that is centred around the $[M - 8H]^{8-}$ charge state, with the most abundant charged species being $[M - 8H]^{8-}$ and $[M - 9H]^{9-}$, *m/z* 2098 and 1865 respectively. Previous studies have illustrated that two charged states exist for apoCaM under electrospray conditions. The

lower charge envelope seen in this experiment is attributed to the compact, folded protein conformation and is the more abundant species under these physiologically applicable buffer conditions [378]. The higher charge envelope (not seen in this experiment) is likely to be due to the denatured apoCaM conformations [378, 379] and under these conditions is likely to be present only in minor proportions and is detected only under high magnification. The spectrum was transformed to a mass of 16,790 Da, in agreement with the calculated mass (Table 5.6) [166]. This experimentally determined mass confirms the presence of N-terminal acetylation and a trimethyl-Lys residue (Lys115) [328].

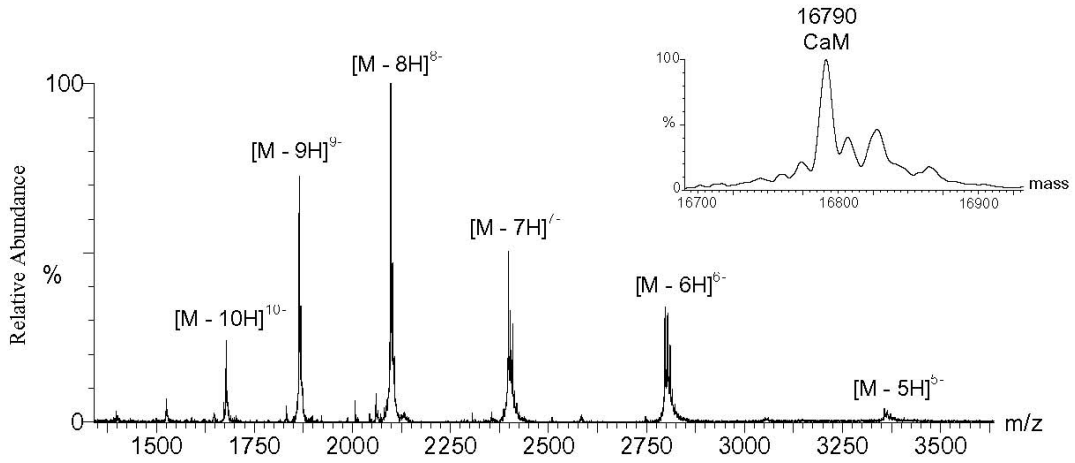


Figure 5.15: ESI-MS spectrum of CaM in the negative ion mode. The inset shows the transformed spectrum.

The negative ion ESI-MS spectrum of CaM obtained with a 8-fold excess of Ca^{2+} is shown in Figure 5.16. There is a charge envelope centred around *m/z* 2419 corresponding to $[M - 7H]^{7-}$. When transformed, this abundant ion corresponds to CaM + $4Ca^{2+}$ ions, with eight protons lost as a consequence of coordination of the four Ca^{2+} ions. With the addition of each Ca^{2+} ion, the transformed spectrum illustrates the addition of 38 mass units to the mass of apoCaM. As the mass of calcium is 40 mass units, the coordination of Ca^{2+} ions is consistent with the loss of two protons from CaM. Hence, the experimental mass of Ca_4CaM is determined as 16,942 Da. This is consistent with previous studies [166, 377]. The change in the charge envelope upon binding of Ca^{2+} ions to apoCaM can be interpreted to reflect the effects of the change in the conformation of the protein upon Ca^{2+} binding [377]. The transformed spectrum also illustrates the additional non-specific Ca^{2+} ion attachment of one or two Ca^{2+} ions. Previous studies have illustrated varied results with binding of up to at least ten Ca^{2+} ions [378]. In contrast, both Loo *et al*. and Veenstra *et al*.

showed the binding of a maximum of four Ca^{2+} ions [166, 380]. Despite this, the results obtained here are consistent with previous studies.

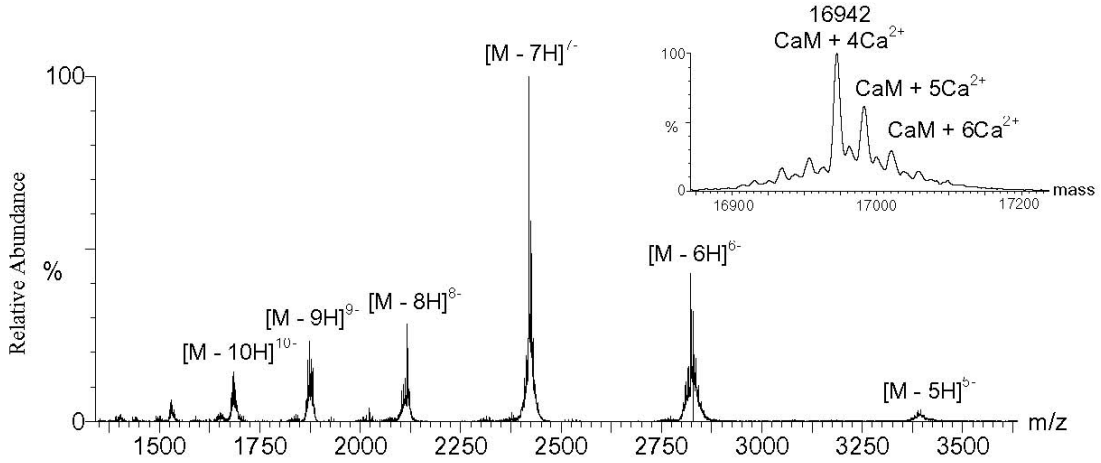


Figure 5.16: ESI-MS spectrum of Ca^{2+} CaM in the negative ion mode. The inset shows the transformed spectrum.

The ESI-MS spectrum of Ca^{2+} CaM with dahlein 5.6 can be seen in Figure 5.17. The most abundant ion present in the spectrum is *m/z* 3188, which corresponds to the $[M - 6H]^{6-}$ charge state of Ca_4CaM-dahlein 5.6 complex. When transformed, a complex mass of 19,135 Da is evident and consistent with the calculated mass (Table 5.6). This spectrum also illustrates the presence of ions for Ca_4CaM, free dahlein 5.6 and apoCaM. Only 1:1 peptide/protein complexes were observed under the conditions used. Increasing the peptide/protein ratio to 10:1 did not change the charge distribution of the spectrum, providing evidence for no binding of multiple peptide molecules to the CaM. Thus, a maximum of one dahlein 5.6 molecule is able to bind to CaM. In addition, it was illustrated that four Ca^{2+} ions were required for the formation of the complexes. This is supported by the transformed spectrum (Figure 5.17, inset). No peptide/protein complex was observed in the absence of calcium acetate (data not shown). In Figure 5.17 the charge state at $[M - 6H]^{6-}$ corresponds to the centre of a charge envelope for the peptide/protein complex. Comparison of this spectrum to that of Ca^{2+} CaM alone (Figure 5.16), indicates that the charge state distribution patterns and relative intensities of the signals are altered upon peptide binding. This suggests that a global conformational change in the protein occurs upon binding to a more compact structure.

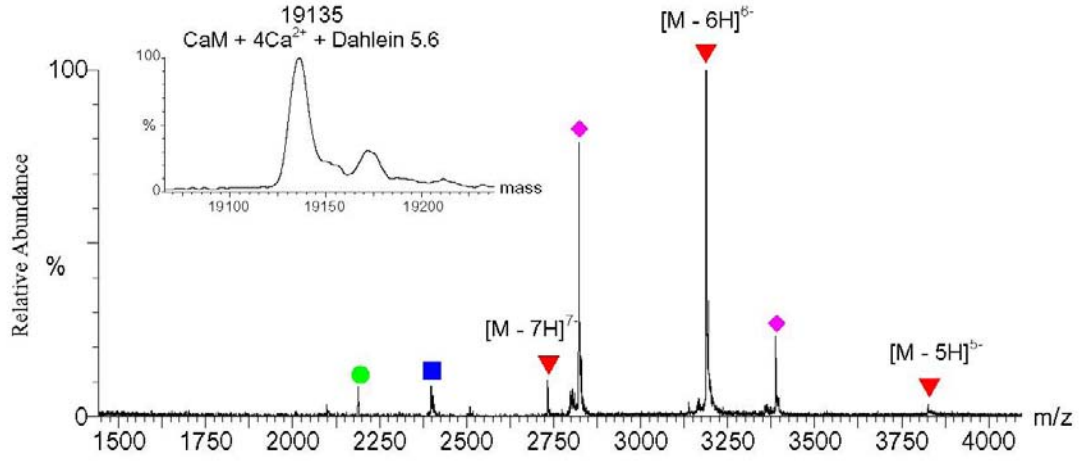


Figure 5.17: ESI-MS spectrum of Ca_4CaM with dahlein 5.6 in the negative ion mode. The inset shows the transformed spectrum. Free dahlein 5.6 (●); apoCaM (■); Ca_4CaM (◆); Ca_4CaM-dahlein 5.6 complexes (▼).

5.2.2.2 ^{15}N HSQC Investigation

Although it is evident that a non-covalent complex exists between dahlein 5.6 and Ca^{2+} CaM (Section 5.2.2.1), these experiments do not provide any detailed information as to the structure and binding interactions of the resulting complex. In order to investigate this further, NMR studies on the structure of the non-covalent complex were undertaken. It was necessary to determine whether the binding of dahlein 5.6 to CaM results in a conformational change in CaM. This was achieved through titration experiments of dahlein 5.6 with ^{15}N-labelled CaM. Unlabelled dahlein 5.6 was added to ^{15}N-labelled Ca^{2+} CaM in increasing concentrations and a high-resolution ^{15}N-^{1}H HSQC spectrum was recorded after each addition of peptide. Chemical shift changes were then tracked by overlaying each of the HSQC spectra (Figure 5.18). Chemical shift changes were considered to be significant if the change was greater than 0.05 ppm in the hydrogen dimension and 0.5 ppm in the nitrogen dimension [381].

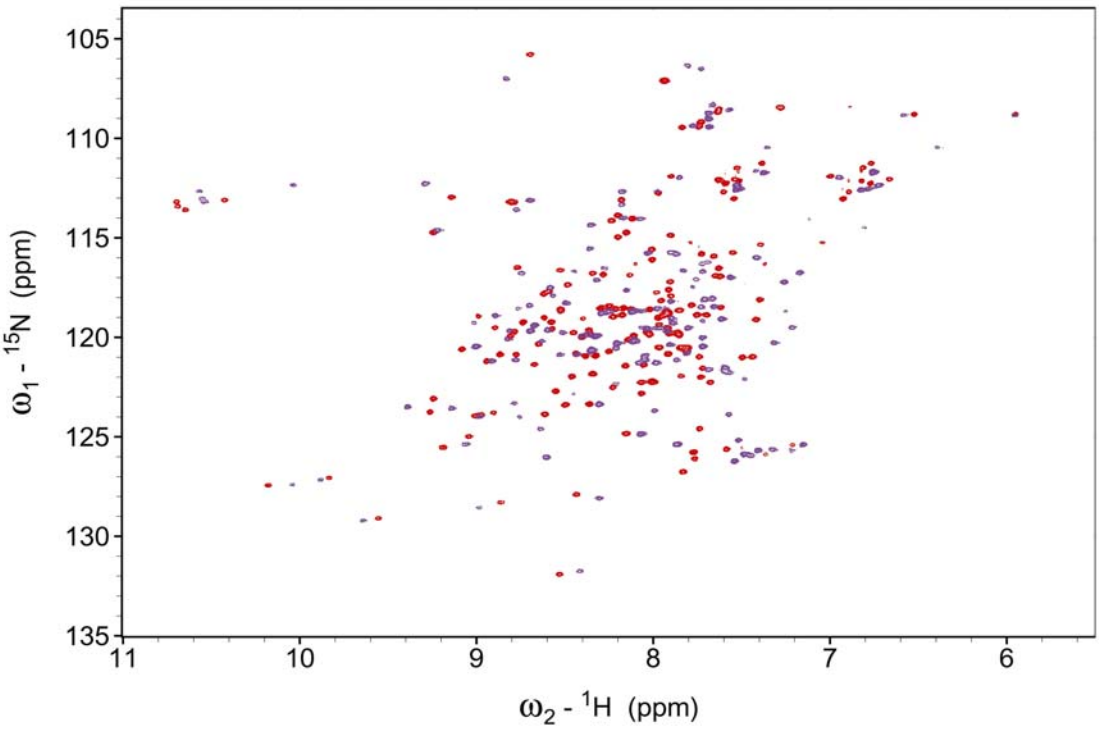


Figure 5.18: ^{15}N-^{1}H HSQC spectra of ^{15}N-labelled Ca^{2+} CaM in the absence of dahlein 5.6 (red) and with the addition of dahlein 5.6 in a 1:1 molar ratio (purple).

Distinct chemical shift changes for a number of residues throughout the Ca^{2+} CaM sequence were observed in the titration series. This suggests that complex formation is evident between Ca^{2+} CaM and dahlein 5.6. Instead of observing chemical shifts changing as a function of concentration, a second set of peaks with distinct chemical shifts appeared after the addition of 0.4 equivalents of peptide (Figure 5.19). The peak intensities of the bound and unbound conformers at 0.6 equivalents of dahlein 5.6 are comparable. Further addition of peptide resulted in a decrease in the unbound peak intensities. This suggests that a 1:1 stoichiometry is present and provides evidence that slow exchange binding is present in the CaM/peptide complex (Figure 5.19). Dahlein 5.6 was completely bound to Ca^{2+} CaM and gave rise to an entirely new set of protein chemical shifts at an equimolar concentration. Addition of the peptide to 4 molar equivalents displayed no further effect on the chemical shifts of the Ca^{2+} CaM suggesting that the protein was fully saturated. The peak intensities for the bound and unbound conformations are approximately equal upon full saturation, indicating that the complex is stable on the NMR time scale. This further supports the evidence that the complex exists in a slow exchange regime.

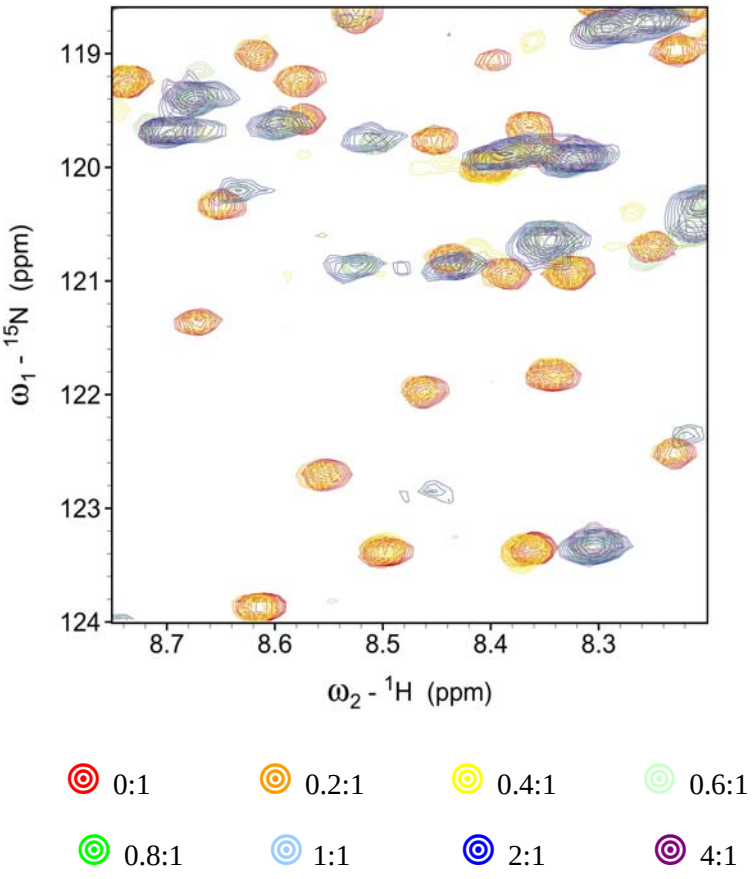


Figure 5.19: Partial overlaid ^{15}N-^{1}H HSQC spectra for the titration of ^{15}N-labelled Ca^{2+} CaM with dahlein 5.6. The peptide/protein molar ratio is indicated in the legend below the spectrum.

Selected cross-peaks in ^{15}N-^{1}H HSQC spectra were assigned on the basis of chemical shifts that were reported previously for unbound Ca^{2+} CaM [382, 383]. No attempt was made to assign either the unbound resonances in the NMR spectra of Ca^{2+} CaM or resonances of the fully bound complex in regions where the concentration of signals would result in ambiguity. Regardless of this, distinct chemical shift changes were seen for a large number of Ca^{2+} CaM resonances. As seen in Figure 5.20, chemical shift changes occurred for selected peaks such as Thr5, Lys21, Gly25, Thr29, Gly61, Lys 94, Gly98, Leu116, Thr117, Asn137 and Ala147. These changes are consistent with a substantial change in both the N- and C-terminal domains of Ca^{2+} CaM conformation following complexation with dahlein 5.6.

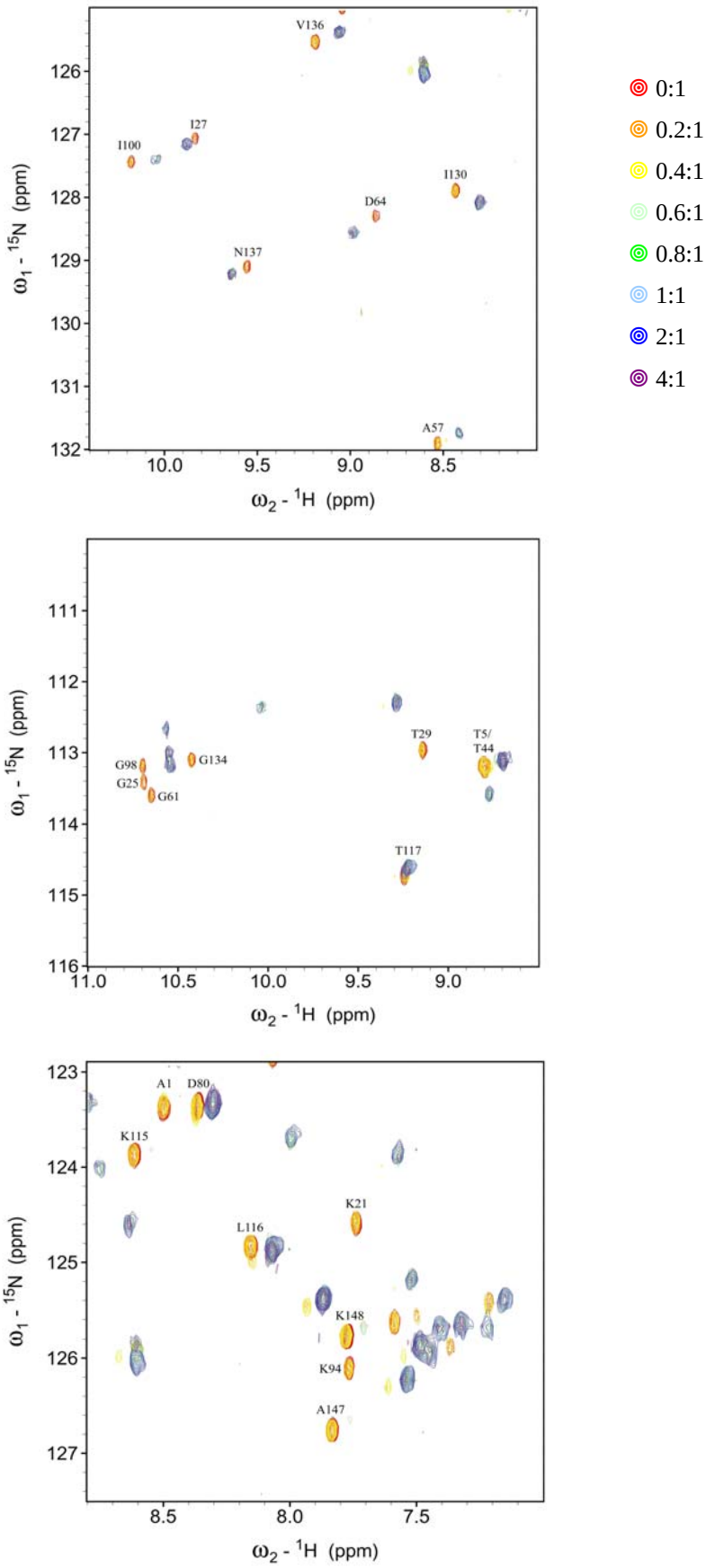


Figure 5.20: Partial overlaid ^{15}N-^{1}H HSQC spectra for the titration of dahlein 5.6 with Ca^{2+} CaM. Selected unbound CaM residues are labelled. The peptide/protein ratio is indicated in the legend.

5.2.3 NMR Spectroscopy of Unbound Dahlein 5.6

Although it is evident that a conformational change occurs in CaM upon binding dahlein 5.6 from the ^{15}N-labelled CaM NMR titrations (Section 5.2.2.2), these experiments do not provide any detailed information as to the structure of the peptide in the resulting complex. In order to investigate this further, additional NMR studies on the structure of the non-covalent complex were undertaken with ^{15}N-labelled peptide.

Commercially synthesised dahlein 5.6 was selectively labelled with ^{15}N at the amide nitrogen of residues Leu3, Gly7, Phe10, Ala15 and Leu18. Labelling at these positions in the peptide allowed the entire length of the peptide to be studied. To enable the structure of a peptide to be investigated by NMR spectroscopy, each resonance needs to be assigned. In aqueous solutions of dahlein 5.6, 2D NMR experiments such as TOCSY, NOESY and COSY experiments were used to assign the proton chemical shifts using the standard sequential method as described by Wüthrich [174] (Section 3.4). Figure 5.21 illustrates the partial TOCSY and NOESY of aqueous dahlein 5.6.

The identification of individual spin systems corresponding to the different amino acid residues present in the peptide sequence was possible in the TOCSY spectrum as the cross-peaks were reasonably dispersed. The COSY spectrum was used to assign the individual resonances in the case of the coincidental amide chemical shifts for Leu2 and Ser5 as well as Ala15 and Lys21 (Figure 5.21). Signal overlap also occurred for Leu6 with Leu2 and Ser5. A combination of COSY spectra and $d_{\alpha N}$ NOE cross-peaks were used to unambiguously assign the resonances.

In the amide region of the NOESY spectrum, no d_{NN} NOE cross-peaks were observed. Sequential assignment was still possible however, due to the presence of αH_i - NH_{i+1} cross-peaks for the majority of the sequential connectivities (Figure 5.21). In the situation where no $d_{\alpha N}$ connectivity was observed, $d_{\beta N}$ NOE cross-peaks provided assistance. As previously stated (Section 5.2.1.1) for Pro20, where no amide proton is present, the connectivities between αH and δH were used to make unambiguous assignment.

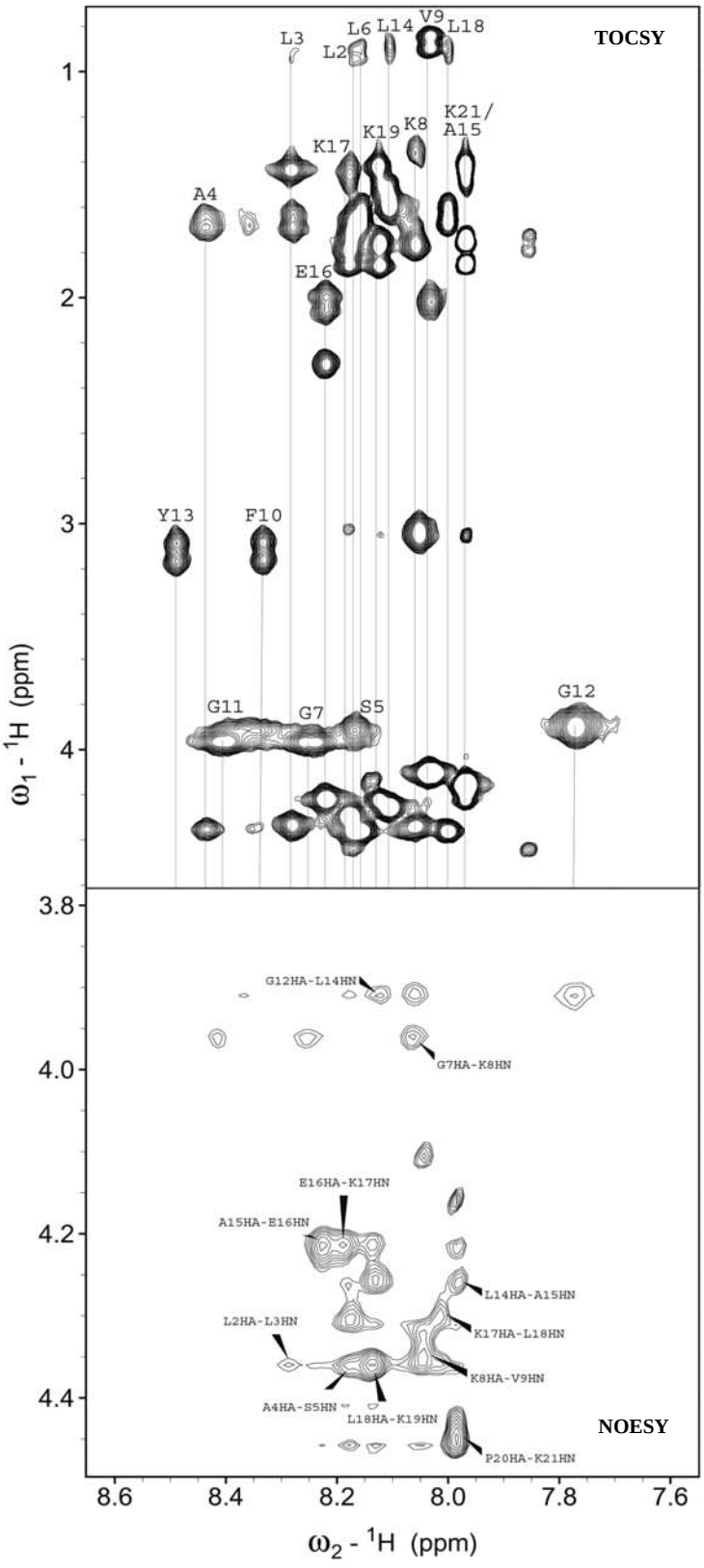


Figure 5.21: Partial TOCSY and NOESY spectra of dahlein 5.6 in H_2O. In the TOCSY spectrum, vertical lines connect resonances in the same spin system. NOEs between sequential αH and amide protons are indicated in the NOESY spectrum.

The αC resonances were assigned from the ^{13}C-^{1}H HSQC spectrum, focusing on the αH ^{1}H/αC ^{13}C region of the spectrum (Figure 5.22). The resonances for Gly1, Leu2 and Lys19 could not be unambiguously assigned due to overlap. A summary of all the ^{1}H and ^{13}C chemical shifts for the aqueous solution of dahlein 5.6 is shown in Table 5.7. In addition, ^{15}N chemical shifts are listed for the labelled residues (Leu3, Gly7, Phe10, Ala15 and Leu18). The assignment of these chemical shifts was achieved using ^{15}N-^{1}H HSQC titration experiments (Figure 5.25). For more details refer to the Section 5.2.4.

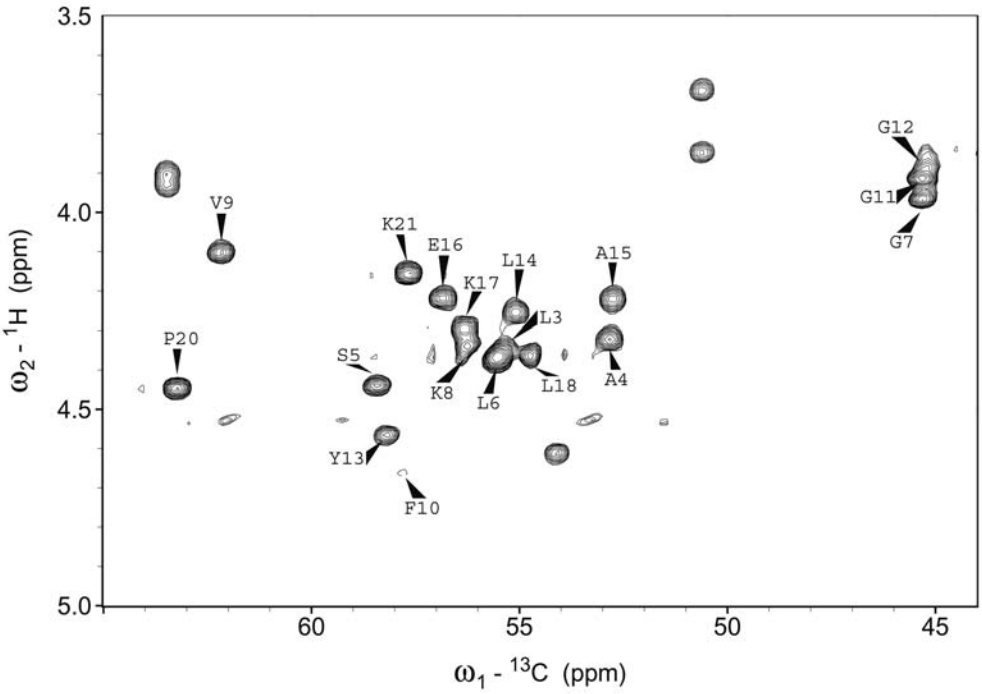


Figure 5.22: αH ^{1}H/αC ^{13}C region of the ^{13}C-^{1}H HQSC for dahlein 5.6 in H_2O. Only the αH-αC correlations are indicated.

Table 5.7: ^{1}H and ^{13}C chemical shifts for dahlein 5.6 in H_2O at pH 6.30, 25°C. n.o. indicates that the resonance was not observed. Amide ^{15}N chemical shifts are also indicated where applicable.

Residue	Chemical Shift					
	HN	**αH**	**βH**	**Other H**	**$\alpha^{13}C$**	**^{15}NH**
Gly1	n.o.	3.99			n.o.	
Leu2	8.17	4.36	1.69	γ-CH 1.64	n.o.	
				δ-CH_3 0.93		
Leu3	8.28	4.34	1.69	γ-CH 1.44	55.32	124.13
				δ-CH_3 0.95		
Ala4	8.44	4.35	1.68		52.84	
Ser5	8.17	4.44	3.92		58.43	
Leu6	8.16	4.36	1.69	γ-CH 1.63	55.53	
				δ-CH_3 0.91		
Gly7	8.25	3.97			45.31	110.47
Lys8	8.06	4.34	1.77	γ-CH_2 1.36	56.25	
				δ-CH_2 n.o.		
				ε-CH_2 3.05		
				ε-NH_3^+ 7.17		
Val9	8.03	4.11	2.02	γ-CH_3 0.90, 0.84	62.17	
Phe10	8.33	4.66	3.16, 3.08	H 2,6 7.30	57.81	125.20
				H 3,5 7.38		
				H 4 7.34		
Gly11	8.41	3.97			45.30	
Gly12	7.77	3.90			45.31	
Tyr13	8.49	4.55	3.17, 3.09	H 2,6 7.16	58.17	
				H 3,5 6.87		
Leu14	8.11	4.26	1.60	γ-CH_2 1.52	55.09	
				δ-CH_3 0.89		
Ala15	7.97	4.22	1.47		52.73	125.60
Qlu16	8.22	4.22	2.06, 2.00	γ-CH_2 2.30	56.55	
Lys17	8.18	4.29	1.86	γ-CH_2 1.46	56.32	
				δ-CH_2 1.79		
				ε-CH_2 3.03		
				ε-NH_3^+ 6.86		

Table 5.7: continues.

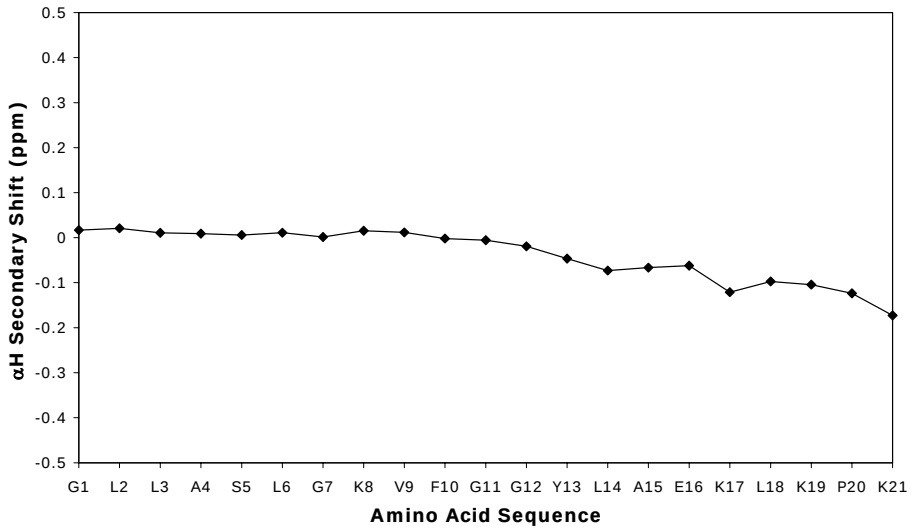
Residue	Chemical Shift					
	HN	**αH**	**βH**	**Other H**	**α^{13}C**	**^{15}NH**
Leu18	8.00	4.36	1.67	γ-CH 1.61	54.74	124.02
				δ-CH_3 0.91		
Lys19	8.12	4.22	1.86	γ-CH_2 1.42	n.o.	
				δ-CH_2 1.77		
				ε-CH_2 3.05		
				ε-NH_3^+ 7.03		
Pro20	-	4.44	2.33, 1.99	γ-CH_2 2.08, 2.03	63.23	
				δ-CH_2 3.86, 3.69		
Lys21	7.97	4.16	1.86	γ-CH_2 1.41	57.68	
				δ-CH_2 1.75		
				ε-CH_2 3.05		
				ε-NH_3^+ 7.28		

5.2.3.1 Secondary Structure of Unbound Dahlein 5.6

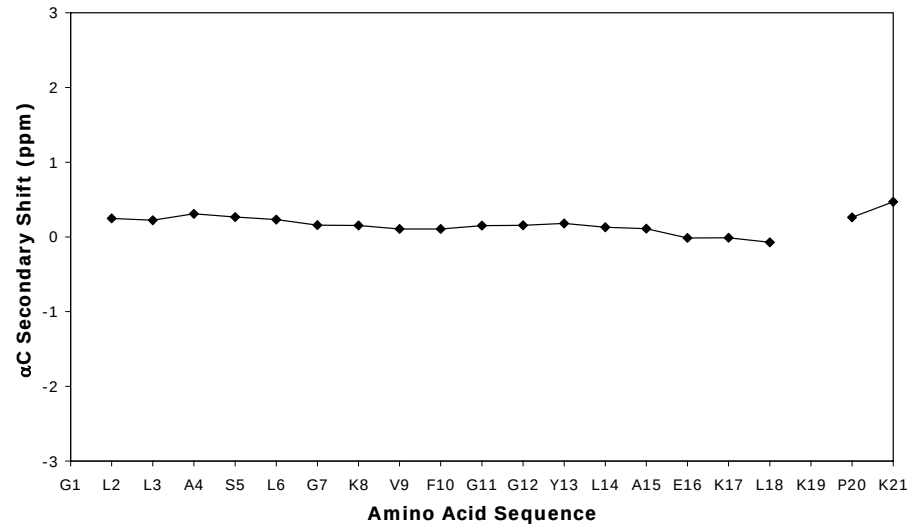
Secondary structural features of a peptide can be identified and characterised using Δδ and NOE connectivity information, without the need to complete extensive structural calculations. The Δδ and NOE connectivity patterns were used to provide information about the secondary structure of dahlein 5.6 in aqueous environments.

The Δδ of a residue describes the variation of the observed chemical shifts for those residues from their corresponding random coil value. Random coil values are obtained from the literature and are determined in water [190]. The Δδ for αH ^{1}H and αC ^{13}C are smoothed over ± 2 residues, whilst the amide ^{1}H values remain unsmoothed (Section 3.6). Negative values indicate that there is an upfield shift from the random coil values, while positive values indicate downfield shifts. The Δδ for aqueous dahlein 5.6 are shown in Figure 5.23.

(a)



(b)



(c)

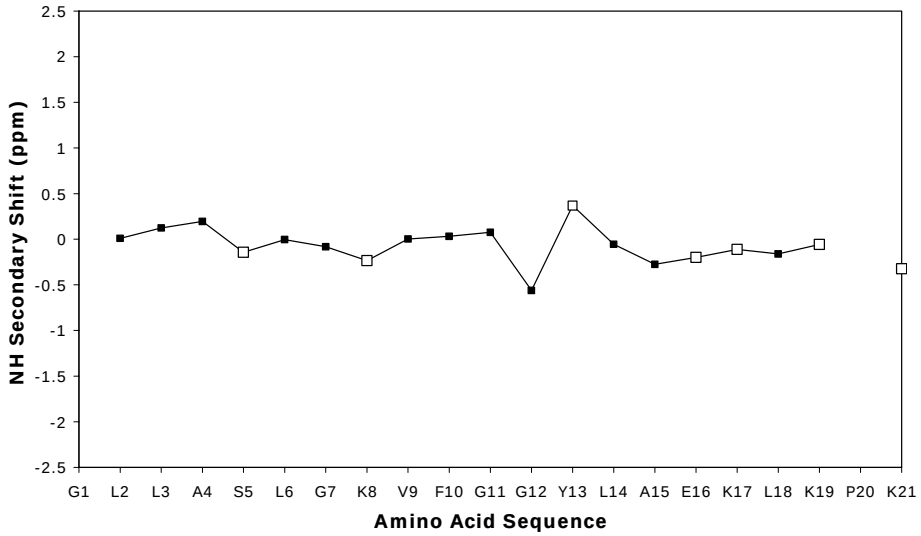


Figure 5.23: (a) αH ^{1}H and **(b)** αC ^{13}C secondary shifts of dahlein 5.6 in H_2O, smoothed over n = ± 2 residues. **(c)** NH ^{1}H secondary shifts, unsmoothed. Negative values indicate an upfield shift from random coil values and positive values indicate a downfield shift. Open symbols represent hydrophilic residues.

The Δδ plots for αH ^{1}H and αC ^{13}C (Figure 5.23) illustrate that there is little variation from the random coil values for the residues. Typically, αH Δδ values of approximately 0.39 ppm upfield are indicative of α helices, whilst shifts of approximately 0.37 ppm downfield are indicative of β sheets or extended conformers [193, 247]. As seen here, the Δδ values approach zero and the smoothed variation is less than 0.1 ppm. Therefore, these shifts can be characterised as random and it can be concluded that dahlein 5.6 exists with conformational flexibility in aqueous environments and it adopts little or no ordered conformation.

Figure 5.24 illustrates the diagnostic NOE connectivities obtained from the NOESY spectrum of aqueous dahlein 5.6. Characteristic patterns of these NOE connectivities can provide information about regions of regular secondary structure. Typically, in regions of α helical structure, strong d_{NN}, medium $d_{\alpha N}$, $d_{\alpha N(i,i+3)}$, $d_{\alpha\beta(i,i+3)}$ and weak $d_{\alpha N(i,i+4)}$ connectivities are observed, however in regions of β sheet structure, strong $d_{\alpha N}$ and weaker d_{NN} connectivities are seen [174, 179, 182, 186]. In this case, many of the characteristic connectivities are absent and the NOE connectivity pattern does not resemble either of these patterns. The main inter-residue cross-peaks observed are the sequential $d_{\alpha N}$, which are present for the majority of the residues and are characteristic of an extended conformation. Relatively few other sequential and medium range NOEs were seen and many of these are ambiguous. Thus, this NOE pattern is consistent with dahlein 5.6 adopting a random or extended conformer in aqueous solution.

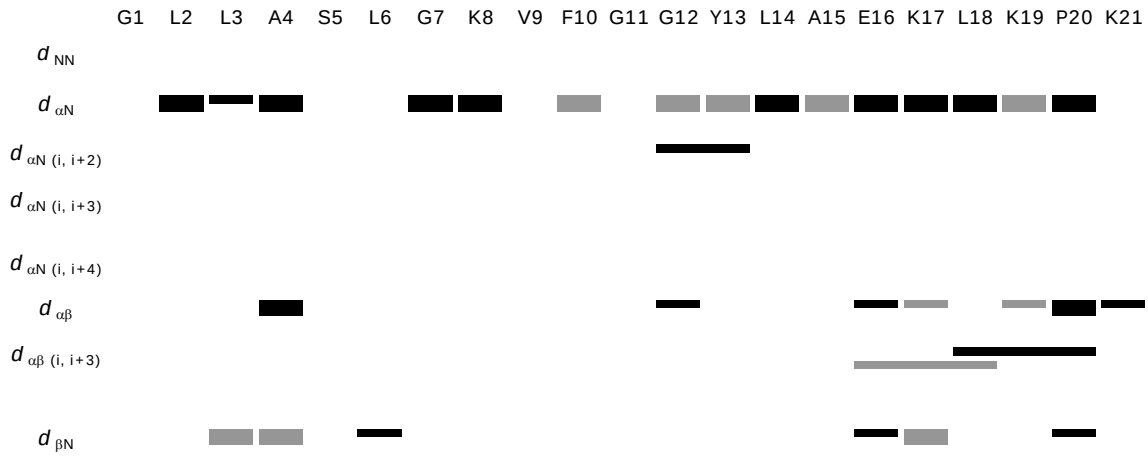


Figure 5.24: A summary of the diagnostic NOEs observed for dahlein 5.6 in H_2O. For Pro where no amide proton is present, NOEs to δH are shown. The thickness of the bar indicates the relative strength of the NOE (strong < 3.1 Å, medium 3.1 - 3.7 Å, weak > 3.7 Å). Grey shaded bars represent ambiguous NOEs.

From the observations obtained from the Δδ plots and NOE connectivity pattern, it can be concluded that dahlein 5.6 has no defined secondary structure about the majority of its length in aqueous environments. This result is considered to be typical of linear peptides in aqueous solutions [384].

5.2.4 ^{15}N HSQC Titration

Further characterisation of the interaction between dahlein 5.6 and CaM was monitored by ^{15}N-^{1}H HSQC titrations of CaM with ^{15}N-labelled dahlein 5.6. As previously mentioned (Section 5.2.3), dahlein 5.6 was ^{15}N-labelled at Leu3, Gly7, Phe10, Ala15 and Leu18. Unlabelled CaM was added to ^{15}N-labelled dahlein 5.6 in increasing concentrations until a molar ratio of 1:1 was reached. A high-resolution ^{15}N-^{1}H HSQC spectrum was recorded after each addition of protein. The chemical shift changes were tracked by overlaying each of the HSQC spectra (Figure 5.25). Chemical shift changes were considered significant if the change was greater than 0.05 ppm in the hydrogen dimension and 0.5 ppm in the nitrogen dimension [381]. As expected, five signals were observed for the unbound peptide and resolved in both dimensions. Assignment of the unbound residues was achieved based on the amide proton chemical shifts obtained for the free peptide in aqueous solutions (Section 5.2.3). Gly7 was easily recognised as this residue has the most upfield chemical shift for the nitrogen resonance.

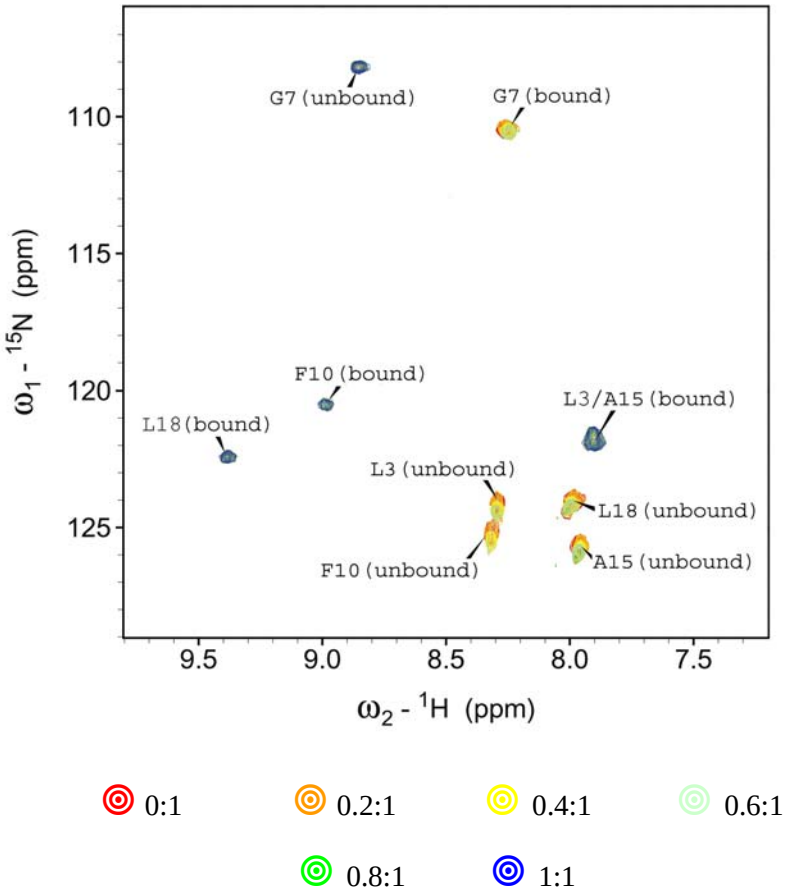


Figure 5.25: Partial overlaid ^{15}N-^{1}H HSQC spectra for the titration of ^{15}N-labelled dahlein 5.6 with Ca_4CaM. The legend is shown below.

A difference in the ^{15}N-^{1}H HSQC spectra was seen after the addition of 0.2 molar equivalents of protein. The chemical shifts did not change as a function of concentration, instead a new set of signals appeared in the spectra, corresponding to the bound state of the peptide. For the bound peptide, only four cross-peaks were observed, with two residues (Leu3 and Ala15) having coincidental chemical shifts in both dimensions. Assignment of these bound cross-peaks will be discussed in Section 5.2.5. Cross-peaks for both bound and unbound residues were present in the spectra at 0.6 molar equivalents. Further addition of CaM reduced the intensity of the unbound cross-peaks providing evidence for slow exchange on the NMR time scale. This also suggests that 1:1 stoichiometry was present for this interaction. A single set of cross-peaks was observed for the bound peptide at equimolar equivalents of Ca^{2+} CaM. The single set of cross-peaks suggests that only one conformation was present for the peptide/protein complex and that a unique binding site was involved in the interaction.

A significant difference in the chemical shift was observed for all five ^{15}N-labelled residues of dahlein 5.6. A change in the chemical shift of a signal indicates a change in the chemical and magnetic environment of the atom, usually associated with a change in the conformational state or complexation of the molecule [181]. This suggests that that peptide has changed its structural conformation upon binding to CaM and that the entire length of dahlein 5.6 is interacting with the protein. The peak shape for the bound residues provides further evidence for this peptide-protein complexation. The bound residues display variations in line width, showing a decrease in intensity and broadening of the cross-peaks (evident in Figure 5.25). This is a result of the bound residues having longer effective correlation times due to an increase in the mass that is associated with complex formation.

5.2.5 NMR Spectroscopy of Bound Dahlein 5.6

The assignment of the resonances for the bound residues in ^{15}N-labelled dahlein 5.6/CaM complex was achieved using heteronuclear NMR experiments. A 2D ^{15}N-edited NOESY spectrum and a 2D ^{15}N-^{1}H HSQC-TOCSY spectrum were acquired for the bound peptide. The assignment of bound Gly7 was easily achieved, as this was the most upfield spin system in the nitrogen dimension. The ^{15}N-^{1}H HSQC-TOCSY spectrum was used to assign the resonances. In the ^{15}N-^{1}H HSQC spectra for ^{15}N-labelled dahlein 5.6 (Figure 5.25), two residues overlapped in both dimensions; Ala15 was found to be coincidental with Leu3.

It was difficult to acquire ^{15}N-^{1}H HSQC-TOCSY spectra that were sensitive enough to all of the correlations in the individual spin systems, despite trying several different spin-lock times. The αH of Phe10 could not be unambiguously assigned due to the presence of the large water resonance. The correlations of several amide hydrogens with αH, δH and γH could not be confidently identified and assigned due to the noise in this spectrum, thus the intra-residue NOE connectivities were used to aid assignment. The ^{15}N-edited NOESY spectrum (Figure 5.26) allowed assignment of the remaining resonances in the bound peptide. Despite being able to assign the residue type to all spin systems, it was difficult to differentiate between Leu3 and Leu18 in the bound peptide. The presence of weak $d_{\alpha N(i,i+4)}$ and $d_{\beta N(i,i+4)}$ cross-peaks corresponding Leu3 connectivities to Gly7, allowed the differentiation to be made between the two Leu residues. Furthermore, the $d_{\alpha N(i,i+3)}$ connectivity between Ala15 and Leu18 confirmed this.

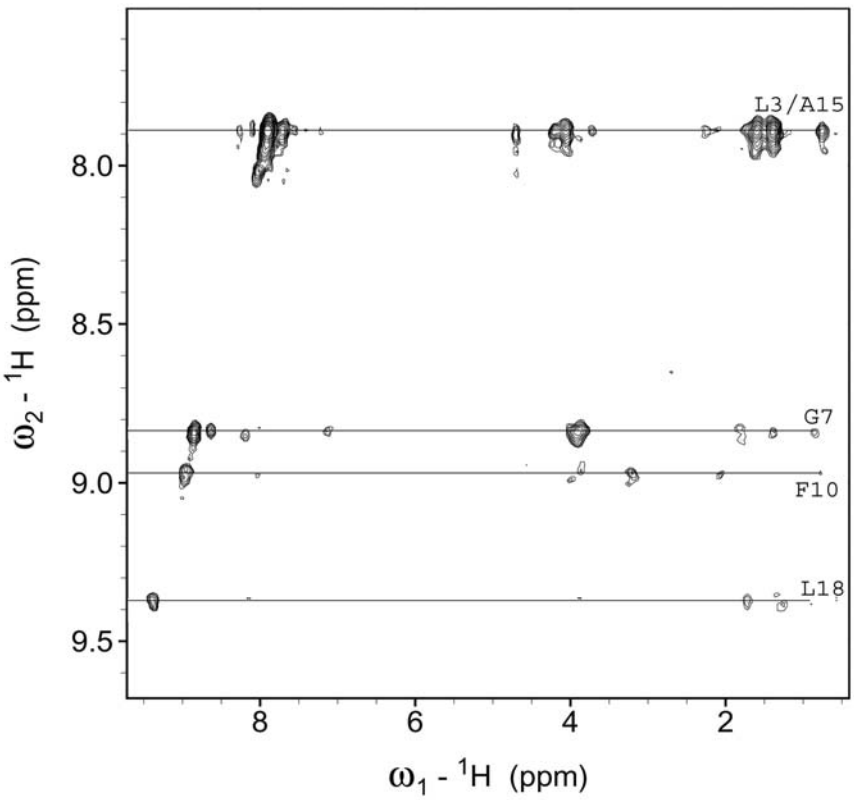


Figure 5.26: Partial 2D ^{15}N-edited NOESY spectrum of ^{15}N-labelled dahlein 5.6 bound to Ca_4CaM in H_2O. Horizontal lines connect resonances of the same residue. NOESY mixing time is 150 ms.

A summary of the ^{1}H and ^{15}N resonances obtained for the ^{15}N-labelled residues of aqueous solutions of bound dahlein 5.6 is shown in Table 5.8.

Table 5.8: ^{1}H and ^{15}N chemical shifts for ^{15}N-labelled residues in dahlein 5.6 bound to Ca_4CaM in H_2O, pH 6.30, 25 °C.

Residue	**Chemical Shift**				
	HN	**αH**	**βH**	**Other H**	**^{15}NH**
Leu3	7.91	4.22	2.18, 1.82	γ-CH 1.60	122.0
				δ-CH_3 0.76	
Gly7	8.86	3.91			108.2
Phe10	8.99	4.55	3.20		120.5
Ala15	7.91	4.06	1.26		122.0
Leu18	9.39	4.19	1.72	γ-CH 1.39	122.4
				δ-CH_3 0.91	

The chemical shift changes for all five ^{15}N-labelled residues were considered to be significant (greater than 0.05 ppm in hydrogen dimension and 0.5 ppm in nitrogen dimension [381]). Differences between the chemical shifts for bound and unbound dahlein 5.6 were particularly apparent for the amide nitrogen and proton, and αH atoms. Changes in these atoms are indicative of a change in conformation of the peptide upon binding to CaM.

Figure 5.27 illustrates the differences in the chemical shifts observed between the unbound and Ca_4CaM bound dahlein 5.6 for Leu3, Gly7, Phe10, Ala15 and Leu18. Upfield shifts are seen for the αH of all five residues. This is thought to be consistent with shifts seen upon the formation of stabilised helical structure [193]. The upfield shifts seen for the ^{15}N resonances, could possibly be associated with a conformational conversion towards a helical secondary structure within the peptide [190, 385]. Variation in the chemical shift of the amide proton resonances has been shown to correlate with variations in the length of its hydrogen bonds [386]. The observed downfield shifts for Gly7, Phe10 and Leu18 ^{1}H resonances may be explained by the formation of short intramolecular hydrogen bonds as the peptide adopts a helical conformation. In contrast, a slight upfield shift was observed for Leu3 and Ala15 amide proton resonances. This may be a result of nearby aromatic residues of the bound CaM inducing ring current effects.

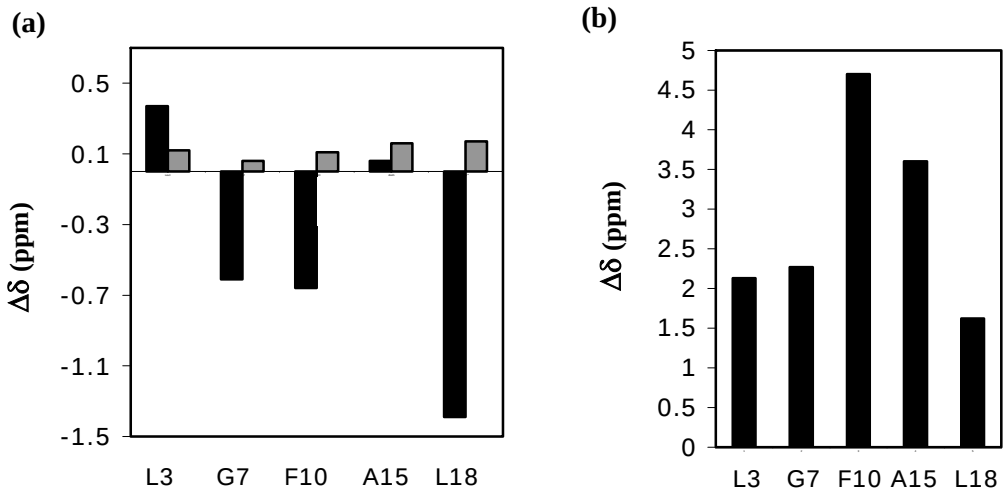


Figure 5.27: The chemical shift changes between unbound and Ca_4CaM bound ^{15}N-labelled dahlein 5.6 for **(a)** NH (black) and αH (grey) proton resonances and **(b)** amide nitrogen ^{15}N resonances of Leu3 Gly7, Phe10, Ala15 and Leu18. Positive values indicate an upfield shift relative to the unbound peptide.

The results presented here suggest that dahlein 5.6 adopts a helical conformation upon binding to Ca_4CaM. This observation is consistent with that of the solution structure of dahlein 5.6 as determined in aqueous TFE and DPC micelles (Section 5.2.1) and most likely represents the bound conformation of the peptide.

5.2.6 nNOS Activities of Dahlein 5.6 Synthetic Modifications

The NMR investigations (Sections 5.2.4 and 5.2.5) indicate that a majority of the dahlein 5.6 peptide is most likely held within the globular Ca^{2+} CaM complex. A number of shorter synthetic modifications of dahlein 5.6 were synthesised in order to determine whether the activity data of these peptides is in accordance with the structural data determined by NMR methods. In addition, the minimum length of the dahlein 5.6 peptide required to maintain the nNOS inhibitory activity could also be investigated. The sequence of the modifications and their nNOS inhibitory activity data are recorded in Table 5.9.

Table 5.9: nNOS inhibitory activities of dahlein 5.6 (from *Litoria dahlii*) and some synthetic modifications.

Peptide	Sequence	MW	IC_{50}		Hill	Charge
			$\mu g.mL^{-1}$	μM	Slope	
Dahlein 5.6	GLLASLGKVFGGYLAEKLKPK-OH	2187	3.5	1.6	2.1	+3
Mod. 1	GLLASLGKVFGGYLAEKLKPK-NH_2	2186	2.6	1.2	2.1	+4
2	GLLASLGKVFGGYLAEK-OH	1737	29.0	16.7	0.8	+1
3	GLLASLGKVFGGYL-OH	1409	57.1	40.5	1.0	+1
4	GLLASLGKVFG-OH	1061	> 700	> 660	-	+1

The most active synthetic modification is the amide modification of dahlein 5.6 (mod. 1). Replacing the C-terminal free acid with an amide increased the positive charge to +4 and lowered the IC_{50} value (from 1.6 to 1.2 μM), thus increasing its nNOS inhibitory activity. The data presented here illustrates that the full sequence of dahlein 5.6 is required for maximum activity. The removal of any residues from the C-terminal end reduces the activity rapidly (Table 5.9). This suggests that the full sequence of dahlein 5.6 is required for maximum binding to Ca^{2+} CaM and supports the finding from the NMR studies that the full length of dahlein 5.6 interacts with Ca^{2+} CaM.

5.3 Discussion

5.3.1 Dahlein 5.6

The results obtained from the solution NMR studies of dahlein 5.6 indicate that the peptide displays α helical secondary structure in the membrane mimicking solvent aqueous TFE and DPC micelles. The resulting structure is likely to represent the solution structure of dahlein 5.6 that persists at the bilayer environment. Dahlein 5.6 displays a more defined structure in DPC micelles. In DPC micelles, dahlein 5.6 exists as a bent amphipathic α helix (Figure 5.13b), with the helix being well-defined over residues 3 – 19. The Schiffer-Edmundson projection illustrating a degree of amphipathic nature in the dahlein 5.6 α helix can be seen in Figure 5.28. The bend in the α helix of dahlein 5.6 occurs at Gly11 and Gly12. This Gly-Gly moiety is commonly associated with bending of α helical secondary structure, such as in splendipherin [78].

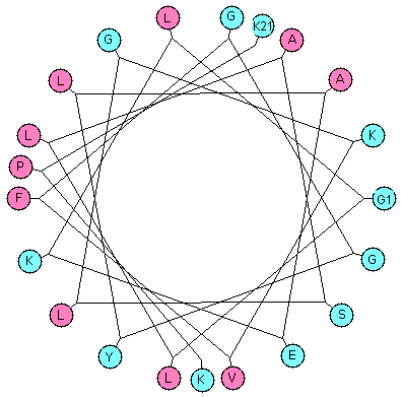


Figure 5.28: The Schiffer-Edmundson helical wheel projection of dahlein 5.6. Hydrophobic residues are illustrated in pink and hydrophilic residues are illustrated in light blue.

5.3.2 Dahlein 5.6 Complexes with Calmodulin

NO is unique among biological signals. Its rapid diffusion, ability to permeate cell membranes and intrinsic instability eliminates the need for extracellular NO receptors or targeted NO degradation. The synthesis, concentration and distribution of NO is controlled by three NOS isoforms that oxidize L-Arg to NO and L-citrulline [297]. The regulatory protein CaM controls the enzyme activity of the nNOS and eNOS isoforms by acting as an

electron shuffler and altering the conformation of the reductase domain to allow the reaction to proceed at the heme site [319]. Numerous amphibian peptides have been identified to inhibit the formation of NO by interacting with Ca^{2+} CaM and preventing its subsequent binding to its corresponding domain in nNOS (Table 5.1).

Dahlein 5.6 has been shown to inhibit the formation of NO by inhibiting nNOS at an IC_{50} value of 1.6 ± 0.3 μM and Hill slope of 2.1. This value is comparable to those of other amphibian peptides that inhibit nNOS and is among the most potent peptide inhibitor of this enzyme that we have isolated to date [14]. Other potent peptide inhibitors include caerin 1.8 [376] and aurein 2.3 [41]. The Hill slope calculated from the concentration response curve for the inhibition of nNOS by dahlein 5.6 is greater than one. This is indicative of positive coordination, suggesting that a non-competitive interaction is causing the inhibition of nNOS. It was proposed that dahlein 5.6, like other amphibian peptides, binds to Ca^{2+} CaM to effect its subsequent inhibition of nNOS.

Previous studies have indicated that peptide binding to Ca^{2+} CaM requires that the peptide: (i) adopt an amphipathic α helical conformation when binding to the protein [352, 353]; (ii) be positively charged [350, 352, 353, 387]; and (iii) display large hydrophobic residues in conserved positions that point to one face in the presumed helical conformation [345]. The proposed extent of the hydrophobic anchoring determines which of the two main binding modes is adopted by the complex (Figure 5.4) [345]. Dahlein 5.6 adopts an amphipathic α helical structure in the membrane mimicking solvent aqueous TFE and DPC micelle (Figure 5.13). Overall, dahlein 5.6 displays a positive charge of +3, with the majority of the charged residues in the C-terminal end of the peptide that would contribute to electrostatic interactions that promote strong binding of dahlein 5.6 to CaM. When the charge of dahlein 5.6 is increased by amidation of the C-terminal, an increase in the inhibitory action of dahlein 5.6 occurs, suggesting stronger binding of the peptide to CaM. On the hydrophobic face of the amphipathic helix of dahlein 5.6, there are four long chain hydrophobic residues (Leu3, Phe10, Leu14 and Leu18) that are available to act as hydrophobic anchors by interacting with Ca^{2+} CaM. The hydrophobic anchors present at both ends of dahlein 5.6 suggest that the peptide will bind to both domains of Ca^{2+}CaM forming a globular, collapsed type conformer [357].

5.3.2.1 Conclusions from Mass Spectrometry Investigations

CaM is an acidic protein with a pI of ~ 4 [377]. Within the pH range used, CaM is negatively charged and binds Ca^{2+} with maximum affinity. Negative ion MS was used to investigate the interactions of dahlein 5.6 with CaM. In this MS mode, the likelihood of observing non-specific metal attachment to CaM is reduced [166]. The Ca^{2+} binding observed here is consistent with previous observations [388]. CaM binding four Ca^{2+} ions was seen, whilst some non-specific binding of a further one or two Ca^{2+} ions also apparent.

The MS studies of dahlein 5.6 and Ca^{2+} CaM indicate that the complex formed has a 1:1:4 dahlein 5.6/CaM/Ca^{2+} stoichiometry. An increase in the peptide/protein ratio to 10:1 did not change the charge distribution of the complex. A comparison of the spectrum of Ca^{2+} CaM alone (Figure 5.16) with that of the dahlein 5.6-Ca^{2+} CaM complex (Figure 5.17) indicate a change in both the charge state distribution patterns and relative intensities of the signals upon the addition of dahlein 5.6. This suggests that a global conformational change occurs in the protein upon binding of dahlein 5.6. Furthermore, four Ca^{2+} ions (two at each end of CaM) are necessary for the formation of the complex, suggesting that the binding of the peptide involves both ends of the Ca^{2+} CaM protein [382, 383].

The absence of dahlein 5.6-Ca_2CaM complexes, in which two Ca^{2+} ions are bound, and the dominance of dahlein 5.6-Ca_4CaM complexes, means that the dahlein 5.6-CaM complexes are only formed when the full complement of four Ca^{2+} ions are present (cf. [331]). Wintrode and Privalov described a peptide binding process to CaM, whereby the first step is either full occupation of the four Ca^{2+} binding sites, or occupation of the two C-terminal domain binding sites of CaM. Subsequently, the peptide will form a bridge between the N- and C-terminal domains of CaM if four Ca^{2+} ions are present. Alternatively, if only two Ca^{2+} ions are bound in the C-terminal domain, the peptide will bind to this domain exclusively [355]. If this view is correct, then it is likely that dahlein 5.6 is interacting with both terminal domains of Ca^{2+} CaM as four Ca^{2+} ions are involved.

Two main binding modes have been identified by NMR studies of specific peptide-Ca^{2+} CaM complexes. Upon complexation with a specific binding peptide, Ca^{2+} CaM undergoes a conformational change to adopt either: (i) a compact, globular shape, with the peptide encompassed in a hydrophobic channel formed by the two terminal domains [345,

349, 353, 389], as that observed by the complex formed between Ca^{2+} CaM and a peptide fragment of MLCK (Figure 5.4a) [353, 389]; or (ii) an extended conformation, where the peptide is bound exclusively to the C-terminal domain [345], as that observed in the complex formed between the binding domain of the plasma membrane Ca^{2+} pump (C20W) and Ca^{2+} CaM (Figure 5.4b) [352].

ESI-MS studies of the 20 residue C20W and Ca^{2+} CaM have indicated that the stoichiometry of this complex is 1:1:2 C20W/CaM/Ca^{2+}. Furthermore, when large excesses of Ca^{2+} ions were added, a complex containing four bound Ca^{2+} ions was observed in the ESI mass spectrum. This is consistent with the binding of C20W lowering the Ca^{2+} affinity of the N-terminal domain so that CaM will binds its full complement of Ca^{2+} ions at higher Ca^{2+} concentrations [390]. This supports the conclusion that dahlein 5.6 is most likely to be interacting with both terminal domains of Ca^{2+} CaM to form a collapsed, globular complex.

5.3.2.2 Conclusions from NMR Investigations

The NMR studies were directed towards the characterisation of the dahlein 5.6 in both its free and CaM complexed states. It is commonly accepted that small linear peptides do not have an ordered secondary structure in aqueous solutions [384]. NMR studies have indicated that dahlein 5.6 exists as a random, extended conformation in water. However, in membrane mimicking solvents, such as aqueous TFE and DPC micelles, dahlein 5.6 adopts a bent amphipathic α helical structure. It is extremely likely that dahlein 5.6 adopts this conformation upon binding to CaM. This is supported by the NMR studies with ^{15}N-labelled dahlein 5.6. The chemical shift changes observed between the unbound and Ca_4CaM bound ^{15}N-labelled residues of dahlein 5.6 are consistent with the formation of a helical secondary structure in the peptide upon binding with Ca^{2+} CaM.

Furthermore, the NMR experiments of ^{15}N-labelled dahlein 5.6 indicate that the labelled region of the peptide is encompassed by the protein. Chemical shift changes are observed between the unbound and bound dahlein 5.6 for all the labelled residues in the region of Leu3 to Leu18, suggesting that these residues are interacting with Ca^{2+} CaM. It was shown that the full length of dahlein 5.6 was required for nNOS inhibitory activity and the removal of residues from the C-terminal end rapidly reduces the inhibitory action. The

minimal length of dahlein 5.6 that is required to maintain inhibitory action was found to be 14 residues, however, this modification was 25 times less active than the native dahlein 5.6 peptide. This implies that the C-terminal end of the peptide is required to increase the binding affinity and strength of dahlein 5.6 to Ca^{2+} CaM, possibly by strengthened electrostatic interactions between the C-terminal Lys residues of dahlein 5.6 and the negatively charged residues of CaM.

Complementary NMR experiments involving the titration of dahlein 5.6 with ^{15}N-labelled CaM, suggest that a more globular conformation of the peptide-protein complex is appropriate (Figure 5.20). Upon binding dahlein 5.6, distinct chemical shift changes occurred throughout the CaM sequence, including the N- and C-terminal domains. This indicates that a substantial change in conformation occurs for the protein upon binding the peptide and the complex that forms has a significantly different structure, instead of mere structural differences at the binding interfaces as expected in the one domain in extended complexes [352]. Both the N- and C-terminal domains of CaM are interacting with the peptide[2] and the protein most likely experiences structural changes towards a collapsed, globular conformation upon binding dahlein 5.6. Analysis of the chemical shift changes for those residues that are readily assigned were found to be consistent with those changes observed for the structure of Ca_4CaM and the peptide fragment of MLCK (Figure 5.4a) [389]. Thus, it is likely that dahlein 5.6, in its helical conformation, is engulfed in the hydrophobic channel of the collapsed-type conformation of the Ca_4CaM complex (as described in Section 5.1.3).

To positively confirm the structure of the dahlein 5.6-Ca_4CaM complex, multi-dimensional heteronuclear NMR experiments are required, with extensive ^{15}N and ^{13}C labeling of both the dahlein 5.6 and CaM. This would provide more information to enable a high-resolution structure of the peptide-protein complex to be determined. Nevertheless, the results presented here provide a good experimental basis for a further understanding of interactions between Ca_4CaM and amphibian nNOS inhibitory peptides.

[2] however, no intermolecular NOEs were identified.

5.4 Experimental

5.4.1 Materials

Dahlein 5.6 and synthetic modifications were synthesised with L-amino acids using the previously reported standard N-α-Foc method [279], by GenScript Corp (Piscataway, New Jersy, USA). All samples used for NMR spectroscopy were shown to be greater than 90 % pure by HPLC and ESI-MS. ^{15}N-L-amino acids were specifically incorporated into the dahlein 5.6 at positions Leu3, Gly7, Phe10, Ala15 and Leu18 by Mimotopes (Clayton, Victoria, Australia). Bovine brain CaM was purchased from Sigma (St. Louis, USA) at greater than 95 % purity and was used as received. ^{15}N-labelled CaM was supplied by Dr. Margit Apponyi from the Department of Chemistry (The University of Adelaide) and based on the method used by Elshort *et al*. [352]. In brief, ^{15}N-labelled CaM was expressed in *E. coli* strain BL21(DE3), using an expression vector pET28 (Novagen, Madison, WI, USA). Expression of CaM was induced by the addition of isopropyl thio-β-D-galactoside (0.1 mM). The cells were then lysed by sonication and the CaM was purified from the supernatant using anion exchange and size exclusion chromatography.

Deuterated d_3-TFE and DPC micelles were purchased from Cambridge Isotope Laboratories (Andover, MA, USA). Water was filtered by a Milli-Q system (Millipore).

5.4.2 Structure Determination by NMR Spectroscopy

5.4.2.1 Sample Preparation

The NMR structure determination of dahlein 5.6 was performed in aqueous d_3-TFE and DPC micelle solutions. Aqueous TFE solutions were prepared by dissolving dahlein 5.6 (4.60 mg, 2.16 x 10^{-6} moles) in a TFE/H_2O (1:1 v/v) mixture to produce a final volume of 0.7 mL. The pH was recorded using a Eutech Cyberscan pH 500 Meter with an AEP 3311 glass-body pH probe (183 x 4 mm, thin stem). A single crystal of DSS was added as a reference.

For the structure determination of dahlein 5.6 in micelles, the peptide (7.45 mg, 3.5 x 10^{-6} moles) and 40 molar equivalents of DPC (53.90 mg, 1.4 x 10^{-4} moles) were dissolved in an aqueous solution containing 10 % D_2O and 50 mM NaH_2PO_4 buffer to produce a final volume of 0.7 mL. The DPC concentration used was in vast excess of the CMC [232]. A peptide/lipid ratio of 1:40 was chosen to give a final solution with approximately one peptide molecule per micelle [234, 236, 237]. The pH was adjusted to 6.0 (using sodium hydroxide) and a single crystal of DSS was added as a reference.

5.4.2.2 NMR Spectroscopy

All NMR experiments were performed on a Varian Inova-600 NMR Spectrometer using a ^{1}H frequency of 600 MHz and a ^{13}C frequency of 150 MHz. Experiments were run at 25 ºC. ^{1}H NMR spectra were referenced to either methylene protons of the residue TFE (3.918 ppm) or to DSS (0.0 ppm). In the ^{13}C-^{1}H HSQC spectra, the ^{13}C F1 dimensions were referenced to TFE $^{13}CH_2$ (60.975 ppm). Additionally, DSS (0.0 ppm) was used.

2D NMR experiments TOCSY, DQF-COSY and NOESY were acquired in phase-sensitive mode, using time proportional phase incrementations in t_1 [280]. Presaturation was used to suppress water resonances. For each experiment, 256 increments, each consisting of 32 transients were acquired over 2048 data points, spread over a spectral width of 6000.2 Hz[1]. NOESY spectra were acquired using a mixing time of 150 ms, while TOCSY pulse sequences included a 60 ms spin-lock. The ^{13}C-^{1}H HSQC experiments were recorded with an interpulse delay of $1/2J_{CH}$ = 3.6 ms, corresponding to J_{CH} = 140 Hz. In the ^{13}C-^{1}H HSQC experiment, 256 increments, comprising of 16 scans were recorded over 4096 data points in the directly detected ^{1}H, F_2 dimension. In the ^{13}C, F_1 dimension, a spectra width of 21117.0 Hz was used. High-resolution 1D ^{1}H NMR spectra were acquired using 0.125 Hz.

For dahlein 5.6 in DPC micelles, all NMR experiments included 200 increments, each consisting of 32 time-averaged scans. The FID in t_2 consisted of 1800 data points spread over a spectral width of 6000.2 Hz[1]. For the TOCSY experiment, a spin-lock of 80 ms was used, whilst the NOESY experiment used a mixing time of 150 ms. For the ^{13}C-^{1}H HSQC

[1] A 90° pulse was calibrated for each experiment and a 1 s recycle delay was incorporated.

experiment, the FID in the F_2 and F_1 dimensions had spectral widths of 6000.2 and 36215.5 Hz respectively.

The resulting 2D NMR experiments were processed using VNMR software (VNMRJ, version 1.1D). Data matrices were multiplied by a Gaussian function in both dimensions prior to zero-filling to 2048 data points before Fourier transformations (^{13}C-^{1}H HSQC: 4096 data points in F_1). The processed 2D NMR matrices contained 2048 x 2048 real data points (^{13}C-^{1}H HSQC: 4096 x 4096).

5.4.2.3 Structure Calculations

Structure calculations were preformed as previously stated in (Section 4.4.8) using Sparky software (version 3.111) to assign the cross-peaks in the TOCSY and NOESY spectra, by way of a standard sequential assignment method [174]. The NOESY cross-peak volumes were converted to distance restraints using a reported method [208]. Where symmetrical pairs of cross-peaks were present, the volume of the larger peak was converted to a distance restraint [205]. $^{3}J_{NH\alpha H}$ values were measured from high-resolution 1D ^{1}H NMR spectra and the corresponding dihedral angles were restrained to: $^{3}J_{NH\alpha H} < 5$ Hz, $\phi = -60^{o} \pm 30^{o}$ and $5 < {}^{3}J_{NH\alpha H} > 6$ Hz, $\phi = -60^{o} \pm 40^{o}$. ϕ angles were not restrained for $^{3}J_{NH\alpha H} > 6$ Hz.

Structures of dahlein 5.6 were generated using ARIA (version 1.2) applied with CNS Solve (version 1.1). Standard ARIA protocol for RMD and SA were used [211], with floating stereospecific assignments employed [281]. Each ARIA run consisted of eight iterations and the structure of the peptides were calculated using standard default ARIA parameters in the run.cns task file. For dahlein 5.6, better convergence was achieved by using the "ambigcutoff" parameters based on those employed by Pari *et al* [282] and by doubling the number of steps in the SA protocol as described by Kang *et al* [220]. For the final iterations, sixty structures were calculated and the twenty structures with the lowest energy were retained for analysis. The final 3D structures were displayed using MOLMOL (version 2K.2) [284] and VMD software (version 1.8.2) [283].

5.4.3 Mass Spectrometry Investigations

MS investigations of CaM, Ca^{2+} and dahlein 5.6 were carried out by Dr. Jennifer Beck and Thitima Urathamakul at the Department of Chemistry, University of Wollongong (Wollongong, New South Wales, Australia).

5.4.3.1 Sample Preparation

Bovine brain CaM (0.3 mg, 1.80 x 10^{-8} moles) was dissolved in 10 mM ammonium acetate (340 μL) and the pH adjusted to 7.6. This sample was treated by dialysis at 4 °C against 2 mM EDTA, 10 mM ammonium acetate solution (4 x 2 L), pH 5.6; 10 mM ammonium acetate (4 x 2 L), pH 5.1; and 10 mM ammonium acetate (2 x 2 L), pH 7.6. CaM concentration was ascertained by measurement of the absorbance at 277 nm and an absorbance coefficient ε_{277} = 3300 $M^{-1}.cm^{-1}$ [391]. CaM concentrations used were typically approximately 100 μM. Ca_4CaM was prepared by adding Ca^{2+} ions as calcium acetate (5 mM in 10 mM ammonium acetate) to produce a 8-fold excess of Ca^{2+} ions over the protein. This concentration was sufficient to observe CaM with full Ca^{2+} binding saturation.

Dahlein 5.6 was dissolved in ammonium acetate (10 mM, pH 7.6) to produce a final concentration range of 5 mM. Small aliquots were added to the CaM solutions to give the following peptide/CaM molar ratios 0.5:1, 1:1, 3:1, 5:1 and 10:1.

5.4.3.2 Mass Spectrometry

CaM, Ca^{2+} CaM and Ca_4CaM-dahlein 5.6 complexes were investigated by negative ion ESI-MS on a Micromass Q-TOF 2 mass spectrometer with a mass range to 10,000 Da and equipped with a Z-spray ionisation source. Samples were injected directly into the Q-TOF 2 using a Harvard model 22 syringe pump (Natick, USA) and a flow rate of 10 $\mu L.min^{-1}$. Spectra were acquired over a mass range of *m/z* 500 – 5000. The following operation parameters were used: capillary voltage of 2.6 kV; cone voltage of 35 V; Ar collision gas energy at approximately 4 eV; and source block and desolvation temperature

of 40 °C. Typically 20 – 30 acquisitions of 5 s were summed to obtain the representative spectra. The data was processed using MassLynx (Micromass, Ltd, version 4.0).

5.4.4 NMR Titrations

5.4.4.1 Sample Preparation

^{15}N-labelled CaM (3.16 mg, 1.89 x 10^{-7} moles) was dissolved in an aqueous solution containing calcium chloride (6.3 mM), potassium chloride (31 mM) and 10 % D_2O at pH 6.3 with a final volume of 0.5 mL. Sodium azide (0.02 %) was added to the aqueous solution as a preservative [382]. Dahlein 5.6 (1.65 mg, 7.56 x 10^{-7} moles), was dissolved in water, the pH adjusted to 6.3 using sodium hydroxide and divided into aliquots so that successive additions would provide the desired peptide/^{15}N-labelled CaM ratios as indicated in Table 5.10. The aliquots were lyophilised, the dried peptide proportions were added to the ^{15}N-labelled CaM aqueous solution in sequence and the pH was adjusted to 6.3 (using sodium hydroxide).

Table 5.10: Dahlein 5.6 quantities used for the NMR titration with ^{15}N-labelled CaM (^{15}N-CaM) at pH 6.3, 25 ºC.

Step	Dahlein 5.6 (mg)	Concentration (mM)	Dahlein 5.6: ^{15}N-CaM
1	0	0	0:1
2	0.083	0.076	0.2:1
3	0.166	0.151	0.4:1
4	0.249	0.227	0.6:1
5	0.332	0.302	0.8:1
6	0.415	0.378	1:1
7	0.827	0.756	2:1
8	1.653	1.512	4:1

^{15}N-labelled dahlein 5.6 (3.29 mg, 1.5 x 10^{-6} moles) was dissolved in an aqueous solution containing calcium chloride (40 mM), potassium chloride (100 mM) and 10 % D_2O at pH 6.3 with a final volume of 0.5 mL, giving a peptide concentration of 3 mM. Sodium azide (0.02 %) was added to the aqueous solution as a preservative [382]. CaM (25 mg, 1.5 x 10^{-6} moles) was dissolved in water and the pH adjusted to 6.3 using sodium hydroxide. The CaM solution was divided into five portions so that successive additions would result in the desired ^{15}N-dahlein 5.6/CaM mole ratio as outlined in Table 5.11. The

aliquots were lyophilised, added to the ^{15}N-dahlein 5.6 solution and the pH adjusted to 6.3 (using sodium hydroxide).

Table 5.11: CaM quantities used for the NMR titration with ^{15}N-labelled dahlein 5.6 (^{15}N-dahlein 5.6) at pH 6.3, 25 ºC.

Titration Step	CaM (mg)	Concentration (mM)	CaM:^{15}N-dahlein 5.6
1	0.0	0	0:1
2	0.5	0.595	0.2:1
3	10.0	1.191	0.4:1
4	15.0	1.787	0.6:1
5	20.0	2.382	0.8:1
6	25.0	2.977	1:1

5.4.4.2 NMR Spectroscopy

All NMR experiments were performed on a Varian Inova-600 NMR Spectrometer using a ^{1}H frequency of 600 MHz and a ^{13}C frequency of 150 MHz. Experiments were run at 25 ºC. ^{1}H NMR spectra were referenced to DSS at 0.0 ppm in the ^{1}H and ^{13}C dimensions while the ^{15}N dimension was centred at 120 ppm.

For NMR titration experiments, the standard gNhsqc pulse sequence in the VNMR library was used. 128 increments, each consisting of 16 transients were acquired over 2048 data points. In the ^{1}H dimension a spectra width of 6000.2 Hz was used whilst in the ^{15}N dimension a spectra width of 2500 Hz was used. All resulting spectra were processed using NMRPipe [392] and viewed using Sparky software (version 3.111).

To investigate the unbound dahlein 5.6, a series of 2D NMR experiments including TOCSY, DQF-COSY and NOESY were performed and acquired in phase-sensitive mode. Presaturation was used to suppress water resonances. Typically, this was achieved by centring the transmitter frequency on the solvent resonances and applying a low power presaturation from the proton transmitter throughout a relaxation delay of 1 s [393]. In the DQF-COSY experiment, solvent suppression was achieved by gradient methods [393]. For each experiment, 200 time increments, each comprising of 128 transients were acquired over 2048 data points, spread over a spectral width of 6000.2 Hz[1]. Mixing times of 100 and

[1] A 90° pulse was calibrated for each experiment and a 1 s recycle delay was incorporated.

150 ms was used for the NOESY and a MLEV-17 spin-lock of 60 ms was included in the TOCSY pulse sequence [185]. The ^{13}C-^{1}H HSQC experiment was recorded with an inter-pulse delay of $1/2J_{CH}$ = 3.6 ms, corresponding to J_{CH} = 140 Hz. In this experiment, 200 increments, comprising of 256 scans were acquired over 2048 data points in the directly detected ^{1}H, F_2 dimension. A spectral width of 21114.8 Hz was used in the ^{13}C, F_1 dimension.

The resulting spectra were processed using VNMR software (VNMRJ, version 1.1D) and data matrices were multiplied by a Gaussian function in both dimensions. This was then zero-filled to 2048 data points before Fourier transformations.

For investigations of the bound dahlein 5.6, ^{15}N-filtered TOCSY experiments were recorded based on the 3D gNhsqctocsy pulse sequence available in the VNMR library. Only the F_2F_3 plane was recorded. For this plane, 64 increments consisting of 256 scans were recorded over 2048 data points, spread over spectral widths of 6000.2 and 2500 Hz in the ^{1}H and ^{15}N dimensions respectively. Spin-locks of 50, 55 and 70 ms were included in the pulse sequence. A ^{15}N-edited NOESY experiment was acquired using the VNMR library 3D gnoesyNhsqc pulse sequence. To achieve the ^{15}N-editing, only one increment in the ^{15}N dimension was recorded. In the ^{1}H dimension, 64 increments, consisting of 64 scans were recorded over 2048 data points, spread over a spectral width of 6000.2 Hz. Mixing times of 150 and 250 ms were used. The resulting filtered spectra were processed using VNMR software (VNMRJ, version 1.1D). For the filtered TOCSY experiment, the F_2F_3 plane was selectively processed, effectively resulting in a 2D spectrum. The data matrices were multiplied by a Gaussian function in both dimensions, prior to zero-filling and Fourier transformation.

5.4.5 nNOS Bioactivity Testing

nNOS inhibition testing was carried out by the Australian Institute of Marine Science (Townsville, Queensland, Australia). Inhibition was measured and analysed by monitoring the nNOS catalysed conversion of [^{3}H]Arg to [^{3}H]citrulline as described by Doyle *et al*. [41].