The three-dimensional structure of caenopore-5 from *Caenorhabditis elegans* determined by means of NMR spectroscopy

A dissertation submitted to the Faculty of Mathematics and Natural Sciences of the Christian Albrechts University in order to acquire a doctoral degree in Biochemistry

by

Justyna Myśliwy

Kiel

April 2007

First reviewer:	Prof. Dr. J. Grötzinger
Second reviewer:	Prof. Dr. M. Leippe
Date of the examination:	30 05 2007
Approved for printing:	30 05 2007

Dean

To my parents

Summary

The caenopore-5 protein encoded by the spp-5 gene is one out of 23 caenopores identified in *C.elegans*. It belongs to the SAPLIP family whose members perform a variety of functions all based on their ability to interact with lipids. It has been shown that caenopore-5 is responsible for the elimination of any alive bacteria present in the digestive tract of the worm. Thus, caenopore-5 seems to contribute to worm nutrition and simultaneously protects against potential pathogens.

The main aim of this project was the determination of the three-dimensional structure of caenopore-5 by means of heteronuclear multidimensional NMR spectroscopy. For that reason, the isotope-labelled ($^{15}N/^{13}C$) caenopore-5 fusion protein was produced in *E.coli*. Protein purification was achieved by a combination of affinity and size-exclusion chromatography and led to high yields of homogeneous protein for NMR experiments. The analysis of the 3D-heteronuclear spectra resulted in the assignment of the amino-acid residues of caenopore-5 to their corresponding resonances. Unexpectedly, the sequential assignment already revealed two conformers of caenopore-5. More than 3000 proton-distance constraints were obtained from the ^{13}C and ^{15}N-edited NOESY spectra, which were used to calculate the three-dimensional structure of caenopore-5. The structure of caenopore-5 displays the saposin-like fold characteristic of the SAPLIP family. It consists of five amphiphatic alpha helices connected by short loops or a kink. They are all arranged in the tertiary structure which resembles a folded leaf stabilised by three disulphide bridges made of six conserved cysteines. Separate structure calculations for the two sets of distance constraints, which correspond to the two different conformers of caenopore-5 revealed that they only differ in the isomerisation of Pro 99. Interestingly, the comparison of the caenopore-5 structure with structures of other members of the SAPLIP family revealed that two non-antimicrobial peptides saposin A and C are the most similar. Furthermore, caenopore-5

showed a remarkable preference for the interaction with negatively charged phosphatidylglycerol liposomes, which served as a model for bacterial membranes. This interaction induced a conformational change in caenopore-5.

Zusammenfassung

Caenopore-5, ist eines von 23 Caenopore-Proteinen, die in *C. elegans* identifiziert wurden, und wird von dem spp-5-Gen kodiert. Es gehört zu der Familie der SAPLIP, welche eine Vielzahl von Funktionen ausüben können. Die meisten dieser Funktionen beruhen auf ihrer Fähigkeit mit Lipiden in Wechselwirkung zu treten. Für Caenopore-5 wurde gezeigt, dass es für die Beseitigung lebender Bakterien im Verdauungstrakt des Wurmes verantwortlich ist. Damit ist es einerseits für die Ernährung des Wurmes von Bedeutung und schützt ihn andererseits gleichzeitig vor potentiellen Pathogenen.

Das Ziel der vorliegenden Arbeit bestand in der Aufklärung der dreidimensionalen Struktur von Caenopore-5 mit Hilfe der heteronuklearen dreidimensionalen NMR-Spektroskopie. Zu diesem Zwecke wurde ein mit $^{15}N/^{13}C$ markiertes Caenopore-5 in *E. coli* hergestellt. Das Protein wurde mit Hilfe einer Affinitäts- und anschließender Größenausschlußchromatographie gereinigt. Die Analyse der heteronuklearen dreidimensionalen NMR-Spektren ermöglichte die Zuordnung der einzelnen Aminosäuren von Caenopore-5 zu den entsprechenden Resonanzen. Bereits aus der sequentiellen Zuordnung konnte überraschenderweise das Vorhandensein von zwei Konformeren abgeleitet werden. Aus den ^{13}C- und ^{15}N-editierten NOESY-Spektren wurden mehr als 3000 Protonen-Abstände identifiziert, welche dann für die Berechnung der dreidimensionalen Struktur von Caenopore-5 verwendet wurden. Die Struktur zeigt die für die Mitglieder der SAPLIP-Familie typische Faltungstopologie. Diese besteht aus fünf amphipathischen α-Helices, welche über kurze *turns* bzw. *loops* miteinander verbunden sind. Die relative Anordnung der Helices zueinander wird auch als *folded leaf* bezeichnet. Diese Anordnung wird durch drei Disulfidbrücken stabilisiert, welche durch die sechs konservierten Cysteinreste der SAPLIP-Familie gebildet werden. Die zwei beobachteten Konformere unterschieden sich lediglich hinsichtlich des

Aminosäurerestes Prolin 99, welcher eine cis/trans-Isomerie aufweist. Ein Vergleich der Caenopore-5-Struktur mit den Strukturen anderer Mitglieder der SAPLIP-Familie ergab interessanterweise die größte Ähnlichkeit zu zwei nicht-antimikrobiellen Peptiden Saposin A und C. Darüber hinaus konnte gezeigt werden, dass Caenopore-5 eine bemerkenswerte Präferenz für die Wechselwirkung mit anionischen Phophatidylglycerol-Liposomen besitzt, welche als Modell für bakterielle Membranen dienen. Die Wechselwirkung mit diesen Liposomen induzierte eine Konformationsänderung des Caenopore-5.

Contents

1 Introduction 1

1.1 Characterisation of *C.elegans* innate immunity 1

1.2 The caenopores of *C.elegans* 2

1.3 Saposin-like protein family 3

1.4 The biological functions of representatives of the SAPLIP family 4

1.5 Introduction to nuclear magnetic resonance spectroscopy 6

1.5.1 NMR spectroscopy of proteins 7

1.5.2 Two-dimensional NMR spectroscopy 9

1.5.3 Three-dimensional NMR spectroscopy 11

1.6 Aim of the study 13

2 Materials and Methods 14

2.1 Transformation of competent *E.coli* strain BL21 pABlac 14

2.2 Protein expression 15

2.3 Protein purification 16

2.3.1 Purification of caenopore-5 from supernatant 16

2.3.2 Refolding and purification of caenopore-5 obtained as inclusion bodies 17

2.4 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) 17

2.5 Size-exclusion chromatography 19

2.6 Protein concentration 20

2.7 Determination of protein concentration 20

2.8 Circular dichroism spectroscopy 21

2.9 Cleavage of caenopore-5 by factor Xa 22

2.10 Effect of different phospholipids on caenopore-5 conformation 22

2.11 Nuclear magnetic resonance spectroscopy 23

2.11.1 Sample preparation for NMR measurements 23

2.11.2 Two-dimensional NMR experiment 23

2.11.3 Three-dimensional heteronuclear NMR experiments 24

2.12 Sequence-specific assignment of NMR signals....24
2.13 Calculation of the caenopore-5 structure....26
2.14 Graphical representation....26

3 Results....27
3.1 Protein expression and purification....27
3.1.2 Purification of caenopore-5 from supernatant....28
3.1.3 Refolding and purification of caenopore-5 from inclusion bodies....28
3.1.4 Size-exclusion chromatography....30
3.2 Circular dichroism spectroscopy....32
3.3 Sequential assignment....33
3.3.1 Sequential assignment of amino-acid residues....34
3.3.2 Assignment of amino-acid side chain resonances....41
3.4 Determination of the secondary and the tertiary structure of caenopore-5....42
3.5 Inspection of residues with double backbone amide chemical shifts....47
3.6 Structure description....50
3.6.1 Conformers of caenopore-5....53
3.6.2 Characterisation of the surface charge distribution....54
3.7 Structural comparison of caenopore-5 with other members of the SAPLIP family....56
3.8 Effect of different phospholipids on caenopore-5 conformation....59

4 Discussion....61
5 References....68
6 Abbreviations....81
7 Acknowledgements....85

1 Introduction

1.1 Characterisation of *C.elegans* innate immunity

Caenorhabditis elegans is a free-living soil inhabitant feeding on microorganism-rich organic material. It belongs to phylum of Nematoda, which are non-segmentated, bilaterally symmetric roundworms. The worm body is composed of two concentric tubes, separated by a fluid-filled space termed pseudocoelom. The outer tube consists of body musculature, gonades, nervous and secretory system and a thin layer of hypodermis, which secretes a collagenous cuticle that encloses the worm body. The inner tube houses the intestine tract beginning with the muscular pharynx armed with a tri-lobed grinder, which crushes bacteria and pumps them for digestion [1].

Since Sydney Brenner introduced *C.elegans* as a model to study animal development and behaviour, it is has become an interesting tool which is applied in many biological investigations. Recently, *C.elegans* has also been included in the research on immunity, as the soil dweller, being in continuous contact with microbes, evolved protective responses against pathogens.

C.elegans utilises a diversity of defence mechanisms which involve exclusively the innate immune system. The physical barriers like a thick and tough cuticle and chitinous membrane lining the intestinal epithelium provide protection of the most exposed surfaces of the worm against bacterial infections [2], [3]. Furthermore, the worm exhibits behavioural reactions associated with recognition and evasion of potential pathogens [4]. However, a number of studies on the immunity of *C.elegans* revealed the evolutionally conserved signalling pathways and many effector molecules involved in pathogen elimination. It is assumed that the following effector molecules target microbes in *C.elegans*: lysozymes, lipase, ABF (antibacterial factor) peptides, glycine and tyrosine-rich peptides and caenopores [5], [6], [7], [8], [9]. In addition, some of them were shown to be under the control of three distinct but undoubtedly intersecting signalling pathways: the TGF-β-like pathway, the insulin-like

receptor pathway and the mitogen-activated protein kinase (MAPK) pathway [10], [11], [12]. Nevertheless, the molecules involved in direct pathogen-associated molecular patterns (PAMPs) recognition have not yet been identified. The TOL-1 receptor, which is the only representative of Toll-like receptors in *C.elegans* and C-type lectins represented by 125 members are likely to contribute to PAMPs recognition, but the corresponding pathogen components are not known [5], [4]. It is also worth noting that the *C.elegans* genome contains a thioester-containing protein (TEP) which has significant sequence similarity to the vertebrate component of the complement system, C3, but its role in the worm's immunity has not yet been shown [13], [14].

Furthermore, a broad range of pathogens, that can infect the worm was identified and the virulence mechanisms which affect pathogenicity in different hosts were studied using *C.elegans* as a model organism [15], [16], [17].

Further detailed studies on the worm's immunity can identify novel factors along with their function and help to elucidate the molecular mechanisms underlying the host-pathogen interactions.

1.2 The caenopores of *C.elegans*

In the fully sequenced genome of *C.elegans,* a group of 23 spp genes was identified. The information about their sequences can be found in a data base on genetics of members of phylum Nematoda: www.wormbase.org. The peptides encoded by spp genes are termed caenopores and consist of a short extension of hydrophobic amino acids representing a signal peptide and a single domain with homology to the peptides of the saposin-like protein family (saposin-like domain) [8], [9]. The knowledge about their biological role is still vague. Only some functional data were provided for a few of them. The product of the recombinantly expressed spp-1 gene was the first caenopore for which the antibacterial function *in vitro* was demonstrated [8]. Furthermore, the spp-1 gene along with the spp–12 gene were upregulated by the FOXO–like transcription factor DAF-16, which is involved in lifespan extension and significant resistance to infections [11].

The expression of two other genes, spp-3 and spp-18, was induced upon *P.aeruginosa* infection suggesting their role in antibacterial defence [18].

Recently, RNAi-mediated gene-silencing of several spp genes has revealed a very interesting phenotype of the worm with a loss of the spp-5 gene function. The *C.elegans* with a functional knockdown of the spp-5 gene could hardly survived on *E.coli* lawns and the intact bacteria were present inside its intestine tract [9]. The spp-5 gene (T08A9.9 gene, GenBank U40417), which encodes caenopore-5, is expressed only in the cells of the intestine epithelium. In addition, recombinant caenopore-5 was shown to possess antimicrobial activity *in vitro* [19].

This data support an involvement of caenopores in the defence against invading pathogens with particular importance of the antimicrobial activity of caenopore-5 for *C.elegans* survival and growth on bacteria which are also the worm's main source of nourishment.

1.3 Saposin-like protein family

The saposin-like protein (SAPLIP) family is a large family of proteins (about 80 amino acids), which is named according to saposins that are small heat-stable, glycosylated peptides located in lysosomes [20], [21], [22]. More than 200 representatives of this family have been identified until now. In case of many of them only gene sequences are known, but their biological functions have still not been established [23]. They are widely spread in very distant evolutionary organisms like the amoebapores of the parasitic protozoan *Entamoeba histolytica* [24], [25], plant-specific insert (PSI) with homology to the saposin-like domain present in precursors of plant aspartic proteinases [26], NK-lysin and granulysin expressed by porcine and human natural killer (NK) and cytotoxic T cells [27], [28]. The comparison of sequences of members of the SAPLIP family shows high diversity. Nevertheless, some conservative features are distinguished, mostly the distribution of hydrophobic amino acids and six cysteines residues, forming three disulphide bridges. In spite of differences on the primary level of protein organisation, the characteristic feature of all representatives of this family is their fold, termed saposin-like fold.

It is described as a bundle of four or five alpha helices arranged in a characteristic tertiary structure, mainly stabilised by three disulphide bridges [29]. Noteworthy are two bacterial peptides: bacteriocin AS-48 from *Enterococcus faecalis* and the PAS factor from *Vibrio vulnificus*, which lack sequence homology. However, only on the basis of their structural similarity are they considered to share a saposin-like fold with members of the SAPLIP family [30], [31].

1.4 The biological functions of representatives of the SAPLIP family

The members of the SAPLIP family perform a wide variety of their biological tasks as autonomous domains or as parts of multidomain proteins.

They participate in enzymatic reactions like the four saposins A, B, C and D, which are cofactors of sphingolipid hydrolysis in lysosomes [32], [33], [20], [34]. The exact mechanism of their action is still under debate, but mostly they are thought to interact with lipids in the lysosomal membrane. They bind to sphingolipids and induce membrane perturbation (saposin C and saposin D) [35], [36] or extract them from membranes and form soluble protein-lipid complexes (saposin B and saposin A) facilitating their degradation by hydrolases [37], [38].

Multidomain enzymes like acyloxyacyl hydrolase (AOAH), acid sphingomyelinase (ASM) and plant aspartic proteinases, prophytepsin and procardosin A, possess saposin-like domain [39], [40], [41], [42]. The role of the saposin-like domain in these enzymes is diverse. It is important for substrate recognition and the enzyme's activity in both AOAH and ASM [43]. [44], [40]. In plant aspartic proteinases PSI seems to be mostly responsible for vacuolar targeting of an immature enzyme to vacuoles, where it is proteolytically cleaved [45], [46], [47]. It is worth mentioning that all plant aspartic proteinases have an unusual PSI with interchanged amino and carboxy-terminal fragments of the saposin-like domain, called swaposin [48].

The interaction of members of the SAPLIP family with a wide variety of lipids can contribute to their distribution and organisation in cellular membranes.

The most prominent example is the surfactant protein B (SP-B), which is indispensable for the reduction of the surface tension by the organisation and stabilisation of the phospholipid surface in lung alveoli [49], [50], [51].

For several members of the SAPLIP family many studies were published on cytolytic activity against mammalian cells and a wide spectrum of microorganisms [52], [53], [54]. They are powerful effector molecules in the pathogenicity of some organisms like the amoebapores [24], naegleriapores from *Naegleria fowleri* [55], saposin like-proteins from *Fasciola hepatica* [56], clonorin from *Clonorchis sinensis* [57]. On the other hand, they play a significant role in the innate immune system as, for example porcine NK-lysin and human granulysin, which are expressed in antigen-activated cytotoxic T-lymphocytes and NK cells to participate in the lysis of virally and bacterially infected cells, microbes and tumors cells [27], [58], [59], [52]. The biochemical and structural data revealed that the cytolytic peptides exert their lytic activity in different ways. They either generate stable pores in membranes (amoebapores) or destabilise the membrane by forming a carpet-like layer (granulysin and NK-lysin) [60], [30]. Notable are recently published hemolytic and antimicrobial activities of the surfactant protein B *in vitro* against both Gram-positive and Gram-negative bacteria, which are negatively modulated by its binding to alveolar phospholipids *in vivo* [61]. Furthermore, PSI of plant aspartic proteinase, cardosin A, was shown to interact with phospholipid membranes and induce vesicle leakage [62]. This result suggests that the proteolytically separated swaposin domain might act in the defence against pathogens in plant cells.

However, representatives of the SAPLIP family are not only involved in immunity by means of their cytolytic action. Recent studies have shown that saposin C and saposin A play an indispensable role in loading lipid antigens onto the antigen presenting molecules CD1b and CD1d of T-lymphocytes [63], [64].

1.5 Introduction to nuclear magnetic resonance spectroscopy

The phenomenon of the nuclear magnetic resonance (NMR) was first described independently by Bloch and Purcell [65], [66]. It arises from the fact that the nuclei of some elements have the property of angular momentum termed spin. The spin angular momentum *S* is defined by the equation:

$$S = h \cdot I \cdot (I+1) / 2\pi$$

where *h* is the Planck constant ($6.62 \cdot 10^{-34}$ Js), *I* is the spin quantum number which is characteristic for a given nucleus and takes the values *I* = 0, ½, 1, 3/2, 2, 5/2,...

Nuclei with a spin angular momentum produce a magnetic field described by a nuclear magnetic moment μ:

$$\mu = \gamma \cdot S$$

where γ is the gyromagnetic ratio, characteristic for a given type of nucleus. All nuclei with a spin quantum number $I \neq 0$ have a magnetic moment, but for NMR spectroscopy only nuclei with a quantum number $I = 1/2$, like ^{1}H, ^{15}N, or ^{13}C are considered to be of importance [67]. A nucleus with a spin $I = 1/2$ placed in an external static magnetic field B_o adopts two opposite alignments along the magnetic field axis and precesses with respect to this axis (see Figure 1). The frequency of the precession is called Larmor frequency and is proportional to the strength of the external magnetic field and the gyromagnetic constant of a particular type of nucleus:

$$\omega = \gamma \cdot B$$

Due to the energy difference between these two alignment states defined by:

$$\Delta E = \gamma \cdot B \cdot h / 2\pi$$

the spin transition from the lower energy state can be forced by applying a radio frequency (RF) pulse with an energy equal to the difference between these

states, fulfilling the nuclear magnetic resonance condition. The excited nucleus reverts to an equilibrium state after the RF pulse, inducing a measurable signal.

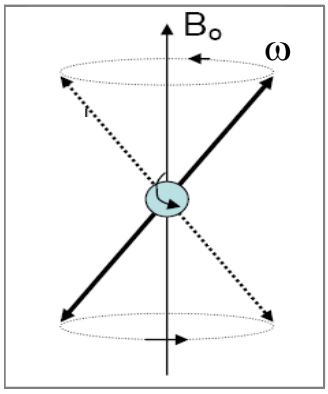


***Figure 1**: A magnetic nucleus precesses with Larmor frequency ω in a static magnetic field B_o.*

1.5.1 NMR spectroscopy of proteins

Nuclear magnetic resonance spectroscopy is a widely used method for three-dimensional structure determination of proteins. It was introduced by Wüthrich for the study of protein structures in solution at atomic resolution [68]. It is particulary applied in the study of protein molecules in a size range between 5 and 35 kDa. The NMR signals in a protein molecule originate from ^{1}H, ^{15}N and ^{13}C spins, where ^{15}N and ^{13}C atoms, due to their low abundance in nature, must be introduced into the protein molecule by recombinant expression of the protein in bacteria growing in a medium supplemented by $(^{15}NH_4)_2SO_4$ or/and ^{13}C-labelled glucose.

A protein sample, placed in a static magnetic field B_o, possesses slightly more spins in the lower energy state, contributing to the bulk magnetisation M_o (see Figure 2). Furthermore, the spins precess with different Larmor frequencies while each one experiences a slightly different strength of B_o, which is the effect of an individual chemical and structural environment inside the protein molecule. Due to these frequency variations, signals from particular atoms are distinguishable, which is the basis for structure determination by

applying the NMR spectroscopy, but it does not support the direct structural information. By manipulating the nature, duration and combination of RF pulses, all spins with different frequencies or the spins within a selected frequency range can be excited at the same time and different molecular properties of the sample can be probed [69], [70]. The acquired signal, termed free induction decay (FID), is represented by Figure 3. It is a time domain signal, consisting of a set of sine (or cosine) waves with different frequencies matching the precession frequencies of the individual spins. It decays over time as the spin system relaxes to the equilibrium state.

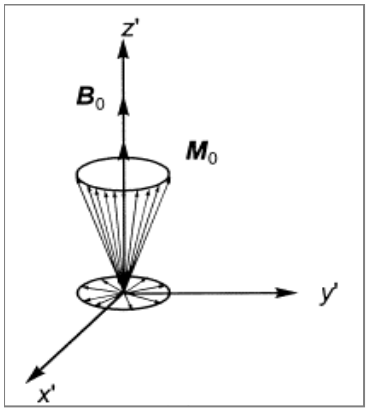


Figure 2: *The bulk magnetisation M_o of a protein sample aligned parallel to the magnetic field B_o under equilibrium conditions, with spins precessing with different Larmor frequencies.*

By means of the Fourier transformation of such an FID, a one-dimensional (1D) frequency spectrum is obtained, which is shown in Figure 3 [69]. The different frequencies are depicted relative to a reference and are termed chemical shifts, which can be described by:

$$\delta = 10^{6} (\nu_0 - \nu_R) / \nu_R$$

The values of the chemical shifts δ are given in ppm (parts per million), where ν_0 and ν_R are the frequencies of a particular spin in the protein sample and a reference, respectively [67]. The introduction of the Fast Fourier Transformation (FFT), established by John Tukey and James Cooley, dramatically changed NMR measurements [71].

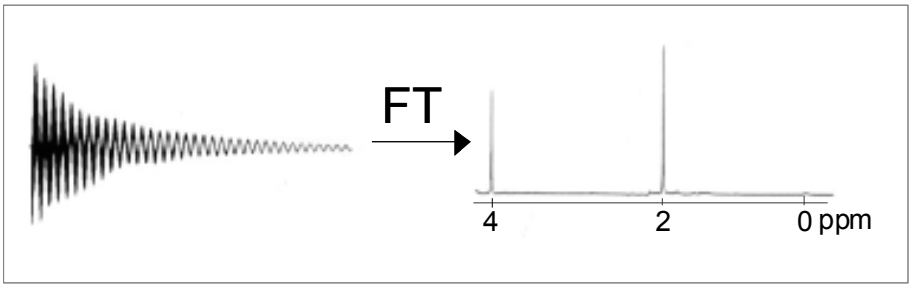


***Figure 3**: The free induction decay (FID signal), which after the Fourier transformation results in the frequency domain spectrum.*

1.5.2 Two-dimensional NMR spectroscopy

A one-dimensional NMR spectrum of a protein with an average size of about 10 kDa contains hundreds of signals which cannot be resolved properly. Therefore, for further detailed analysis, multidimensional NMR spectra are necessary. Not only do they provide an increased resolution by introducing additional dimensions, but furthermore they support the information about interatomic contacts. All of this can be achieved by performing even more complex NMR experiments, consisting of a sequence of various RF pulses separated by delays.

In the two-dimensional (2D) spectrum, which is the basis for generating higher dimensional spectra, the following time periods can be distinguished: preparation, evolution, mixing, and detection time (see Figure 4) [72], [73]. During this sequence of events, the spins precess at their own frequencies, reflecting their chemical environment (evolution time) and interact with one another (mixing period). They interact by two types of magnetisation transfer:

- scalar coupling
- dipolar coupling

The scalar coupling interaction is mediated through electrons. Therefore, only atoms connected through chemical bonds are involved [74]. The dipolar coupling interaction deals with spins, which are in close proximity in space [75].

It involves spins coming from different parts of the protein molecule and not only the sequential neighbourhood.

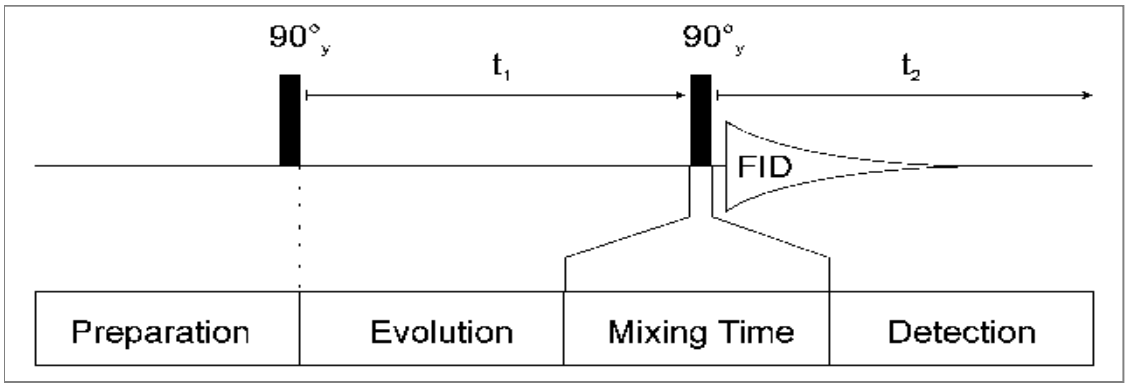


Figure 4: *Scheme of a 2D-NMR experiment in which short time periods are separated by two 90° pulses, and the resultant FID is detected during t_2 [76].*

To obtain a two-dimensional spectrum and decode the information about the spins which interact with each other, the above pulse sequence must be repeated while gradually increasing t_1 and a second Fourier transformation of a set of 1D-spectra with respect to t_1 generates the second dimension (see Figure 5). If two spins interact with each other, the intensity of a signal of one spin is modulated by the frequency of the second one, which can be deciphered in the second dimension.

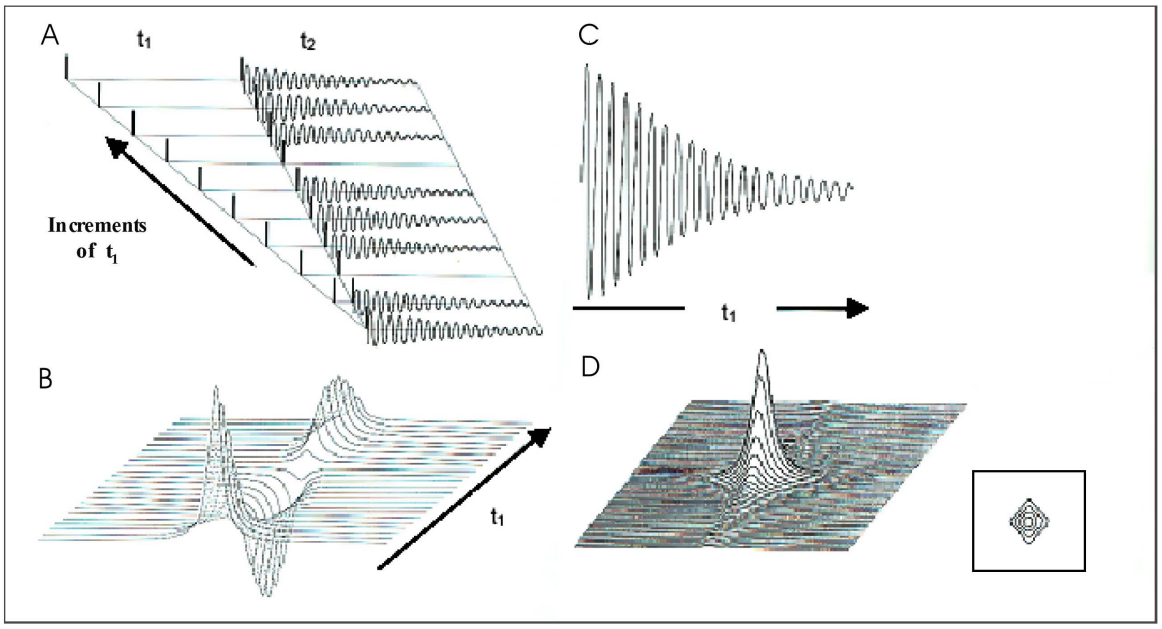


Figure 5: *Schematic representation of a 2D-NMR experiment, where FIDs are collected with regular increments of t_1,* ***(A)****. The Fourier transformation of the FIDs generates frequency spectra with signal intensities modulated in t_1* ***(B)****. A second Fourier transformation of signals along t_1* ***(C)*** *produces a signal positioned in a 2 D spectrum* ***(D)*** *[76].*

1.5.3 Three-dimensional NMR spectroscopy

All three-dimensional (3D) heteronuclear NMR experiments can be classified into double (correlating ^{1}H and ^{13}C or ^{1}H and ^{15}N) and triple-resonance spectra (correlating ^{1}H to ^{13}C and ^{15}N). They provide correlations between protons and heteronuclei exploiting a series of large (one, two or three bonds) heteronuclear scalar couplings to transfer magnetisation. They often correlate backbone amide ^{1}H with its ^{15}N and with alpha ^{13}C or alpha ^{1}H or ^{13}CO or/and side-chain ^{1}H and ^{13}C of its own residue and/or preceding one, providing connections between neighbouring residues. In addition, 3D-heteronuclear NOESY experiments take advantage of the presence of dipolar coupling interactions, supplying data in form of interatomic distances which are the basis for the structure determination [77].

The 3D-NMR experiments are constructed on the basis of two-dimensional ones by introducing an additional evolution time (t_2) and a second mixing period to the pulse sequence. The signal is observed during detection time t_3. It additionally implicates two independent increments according to t_1 and t_2, which are separately Fourier transformed to give a three-dimensional spectrum [78]. They utilise blocks of pulse sequences directed towards the protons and the heteronuclei of interest. The pulse sequences are mostly designed to first excite the protons and then transfer their magnetisation to the coupled heteronuclei, and at the end of the pulse sequence, the magnetisation is transferred back from the heteronuclei to the protons [79], [80]. The purpose of this directed magnetisation transfer is an increase in sensitivity in comparison to other heteronuclei due to the larger gyromagnetic ratio γ of protons. The magnetisation transfer is even more efficient if relatively large scalar couplings are used between protons attached to heteronuclei or between pairs of heteronuclei. Therefore, it takes less time to correlate nuclei and the signal broadening due to relaxation is decreased. Furthermore, decoupling pulses are applied during evolution times to eliminate line splitting of spins caused by scalar couplings and increase both the resolution and sensitivity [81].

1.6 Aim of the study

Caenopore-5 is representative of 23 caenopores identified in *C.elegans*, which on basis of their sequence homology are classified to the SAPLIP family. The knowledge about the biological role of caenopores is still not complete, but they are believed to contribute the innate immunity of the worm. Interestingly the inhibition of spp genes activity showed the indispensable role of caenopore-5 in the elimination of bacteria ingested by the worm [9]. Therefore the determination of the three dimensional structure of caenopore-5 and comparison to the known saposin-like structures will provide useful information about the functional properties and structural aspects of the SAPLIP family.

2 Materials and Methods

2.1 Transformation of competent *E.coli* strain *BL21* pABlac

The expression vector pIVEX2.4a (Roche, Mannheim, Germany) encoding for caenopore-5 was kindly provided by Prof. Dr. M. Leippe. The gene encoding caenopore-5 was cloned into pIVEX2.4a using KspI and PstI cleavage sites to generate a fusion protein with an N-terminal $(His)_6$ tag and a factor Xa cleavage site. The resultant protein sequence of caenopore-5 gene fused with a $(His)_6$ tag is depicted below:

```
                                                  ↓
M  S  G  S  H  H  H  H  H  H  S  S  G  I  E  G  R  G  R  S  A  L
S  C  Q  M  C  E  L  V  V  K  K  Y  E  G  S  A  D  K  D  A  N  V
I  K  K  D  F  D  A  E  C  K  K  L  F  H  T  I  P  F  G  T  R  E
C  D  H  Y  V  N  S  K  V  D  P  I  I  H  E  L  E  G  G  T  A  P
K  D  V  C  T  K  L  N  E  C  P
```

The underlined sequence is the factor Xa recognition region with the cleavage site located downstream of this fragment. The arrow indicates the beginning of the caenopore-5 sequence.

This vector was used to transform the competent *E.coli* bacteria strain BL21 (DE3) pABlac. The suspension consisting of 1μl of vector (100 ng/μl) and 10 μl of competent bacteria was incubated on ice for 30 min, followed by subjecting the bacteria cells to a heat shock procedure for 1 min at a temperature of 42 °C. Afterwards, the bacterial suspension was left to cool down on ice and 30 μl of LB medium without antibiotic were added to let the bacteria propagate for about 1 h at 37 °C. Next the bacteria were spread on an ampicillin (50 μg/ml) and kanamycin-containing (30 μg/ml) agar plate in order to select appropriate clones.

2.2 Protein expression

To produce single-labelled (^{15}N) or double-labelled ($^{15}N/ ^{13}C$) protein, the BL21 pABlac bacteria, containing the vector with caenopore-5 fusion protein, were cultured in minimal medium, either with ^{15}N-labelled ammonium sulphate $(^{15}NH_3)_2 SO_4$ or with ^{15}N-labelled ammonium sulphate $(^{15}NH_3)_2 SO_4$ and ^{13}C-labelled glucose, together with 50 μg/ml ampicillin, in shaker flasks at 37 °C. All isotope-labelled chemical compounds were purchased from Cambridge Isotope Laboratories (Saarbrücken, Germany).

The bacteria growth was monitored by measuring the optical density at 600 nm in an UV–spectrophotometer (U-2001, Hitachi, Japan). After OD_{600}=0.4 was reached, isopropyl-β-D-1-thiogalactopyranoside (IPTG) was added to the culture to a final concentration of 1 mM to induce protein production. The cell growth continued for a further 12-14 hours in order to obtain the protein.

Minimal medium:

2.0 g/l α-D-Glucose-^{13}C

1.0 g/l ^{15}N Ammonium Sulphate

7.5 g/l $Na_2HPO_4 \cdot H_2O$

3.0 g/l KH_2PO_4

2.0 g/l NaCl

0.41 g/l $MgCl_2 \cdot 6H_2O$

2.3 Protein purification

The bacteria were harvested by centrifugation at 8000 g for 20 min and then resuspended in 50 mM Tris-HCl and 150 mM NaCl at pH=7.4. This suspension was sonicated at 40 – 50 % power for 1-1.5 min on ice to disrupt the cell membranes (Sonopuls HD 2200, Bandelin electronic GmbH, Berlin, Germany). After centrifugation of the cell lysate, two fractions were obtained that contained caenopore-5 fusion protein: a supernatant with soluble caenopore-5 and a solid phase containing the protein as inclusion bodies.

2.3.1 Purification of caenopore-5 from supernatant

The supernatant fraction was treated with a glutathione redox system at a final concentration of 10 mM reduced and 2 mM oxidised glutathione (Fluka, Taufkirchen, Germany). The incubation took place overnight at room temperature to allow reshuffling of the disulphide bridges of the soluble caenopore-5 fusion protein. Afterwards, the supernatant was loaded onto a Ni^{2+}-NTA agarose column (Qiagen, Hilden, Germany), equilibrated with 50 mM sodium phosphate, 300 mM NaCl and 10 mM imidazole at pH=8 to bind caenopore-5 via its His-tag. The column matrix was washed with a buffer containing 50 mM sodium phosphate, 300 mM NaCl and 40 mM imidazole at pH=8 to remove unspecifically bound protein. The protein was eluted using 50 mM sodium phosphate, 300 mM NaCl and 250 mM imidazole at pH=8.
The elution fractions were pooled and concentrated (see Section 2.6). The buffer was exchanged for 50 mM sodium phosphate at pH=5.2 by using a NAP-10 column (Amersham Biosciences, Freiburg, Germany).

2.3.2 Refolding and purification of caenopore-5 obtained as inclusion bodies

The inclusion bodies were dissolved in 50 mM Tris-HCl, 100 mM dithiothreitol (DTT) (Roth, Karlsruhe, Germany), 6 M guanidine hydrochloride (Gdn · HCl) (Roche, Mannheim, Germany) at pH=7.4 and incubated for 1 h at room temperature to achieve proper solubilisation. The solution was centrifuged at 14000 g for 15 min to remove all non-dissolved material. The resultant supernatant was subjected to refolding by overnight dialysis against 50 mM Tris-HCl, 150 mM NaCl at pH=7.4 at 4 °C. Precipitated protein was removed by centrifugation at 14000 g for 15 min. The redox system at a final concentration of 10 mM reduced and 2 mM oxidised glutathione was added and incubated overnight at room temperature. This solution was applied onto a Ni^{2+}-NTA agarose column and purified in the same way as the supernatant (see Section 2.3.1).

2.4 SDS-polyacrylamide gel electrophoresis (SDS–PAGE)

SDS-PAGE was used to identify the fractions containing the protein of interest and to monitor protein purity under reducing and non-reducing conditions.

The discontinuous Laemmli system with a stacking and a separating gel was used [82]. For setting up the gel cassettes, a mini slab gel system was employed (Mini- Protean Electrophoresis System, Biorad, USA). The protein samples were dissolved in a loading buffer with or without 2-mercaptoethanol, boiled for 5 min at 90 °C and loaded onto a 15 % polyacrylamide gel. The electrophoresis was run at 200 V at room temperature.The protein bands were fixed to the gel and stained in a Brilliant Blue G 250 (Roth, Karlsruhe, Germany) solution for 40 min. For destaining in order to remove the dye that was not bound to protein molecules a aqueous solution containing 10 % methanol and 10 % acetic acid was used.

Caenopore-5 fusion protein was identified by its size relative to the protein molecular weight marker (Fermentas, St. Leon-Rot, Germany) that was used as a reference.

2x concentrated loading buffer:

20 % glycerol

4 % SDS

1.25 mM TRIS-HCl, pH=6.8

0.002 % Bromophenol Blue

10 % Mercaptoethanol (under reducing conditions)

Stacking gel (7.5 % polyacrylamide):

H_2O	1.23 ml
30 % Acrylamide	0.63 ml
0.5 M Tris-HCl, pH=6.8	0.63 ml
10 % SDS solution	25 µl
10 % Ammonium Peroxydisulphate solution (APS)	25 µl
N, N, N', N'- tetramethylethylenediamine (TEMED)	2.5 µl

Separating gel (15 % polyacrylamide):

H_2O	1.38 ml
30 % Acrylamide	3 ml
1.5 M Tris-HCl, pH=8.8	1.5 ml
10 % SDS solution	60 µl
10 % Ammonium Peroxydisulphate solution (APS)	60 µl
N, N, N', N'- tetramethylethylenediamine (TEMED)	3 µl

Brilliant Blue Solution:

25 % Methanol

10 % Acetic acid

0.1 % Brilliant Blue G 250

Destaining solution:

10 % Methanol

10 % Acetic acid

2.5 Size-exclusion chromatography

Size-exclusion chromatography was applied to further purify and determine the molecular weight of caenopore-5 fusion protein. The soluble and refolded protein both from the supernatant and the inclusion bodies fraction was loaded onto a Hi-Load Superdex 75 prep grade (16/60) column (Amersham Biosciences, Freiburg, Germany), equilibrated with 50 mM sodium phosphate and 5 mM EDTA at pH=5.2. The column was connected to an ÄKTA-FPLC system controlled and supervised by the UNICORN software (Amersham Biosciences, Freiburg, Germany). The elution profile was monitored by measurement of UV-absorbance at 280 nm. Fractions containing caenopore-5 fusion protein were pooled and concentrated (see Section 2.6). The presence of caenopore-5 in these fractions was confirmed by SDS-PAGE.

For calibration of the column a low molecular weight gel filtration calibration kit (Amersham Biosciences, Freiburg, Germany) dissolved in 50 mM sodium phosphate and 5 mM EDTA at pH=5.2 was used. The calibration kit consists of a mixture of proteins with known molecular weights in a range from 6.5 to 75 kDa: conalbumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), ribonuclease A (13.7 kDa) and aprotinin (6.5 kDa).

2.6 Protein concentration

The protein was concentrated using Vivaspin 20 concentrators with a 3000 MWCO polyethersulfone membrane (VivaScience, Hannover, Germany).

2.7 Determination of protein concentration

UV-spectrophotometry was used to determine protein concentrations. Protein-UV spectrophotometry depends on the presence of the phenolic group of tyrosine, the phenyl group of phenylalanine and the indolic group of tryptophan residues in a protein molecule, which are responsible for UV light absorption at 280 nm. The following equation was used for the calculation of the protein amount:

$$c = A_{280} \cdot M / (\varepsilon_{280} \cdot d)$$

where c is the protein concentration in g/l, A_{280} is the absorption at 280 nm, M is the molecular weight in g/mol, $\boldsymbol{\varepsilon_{280}}$ is the absorption coefficient in $M^{-1}cm^{-1}$ and d is the path length of the sample in cm.

The absorption spectra for caenopore-5 solutions were recorded in a range of 240-320 nm in a UV–spectrophotometer (U-2001, Hitachi, Japan). Measurements were made in 1 cm quartz cuvettes, where the buffer was used as the reference. For calculating the caenopore-5 fusion protein concentration the following parameters were used: ε_{280} = 2920 $M^{-1}cm^{-1}$ and M = 10988 g/mol. The extinction coefficient (ε_{280}) was calculated according to the method of Gill and von Hippel [83].

2.8 Circular dichroism spectroscopy

The refolded caenopore-5 fusion protein was investigated by far-UV circular dichroism (CD) spectroscopy. In addition, changes of the CD-signal caused by an increase in temperature were measured to provide estimates of the protein stability.

In the far-UV range, the absorption is principally due to the peptide bond so the different types of regular secondary structure found in proteins give rise to characteristic CD-spectra [84]. The circular dichroism originates from the differential absorption of right- and left-handed circularly polarized components of plane polarized light. The resultant light is elliptically polarized, therefore circular dichroism spectra are presented in terms of the mean residue ellipticity *[θ]* at wavelength *λ* defined by:

$$[\theta]_{\lambda}^{MRW} = MRW\theta_{\lambda} / (10 \cdot d \cdot c)$$

where *MRW* is the mean residue weight for a particular protein (in Da), *θ* is the observed ellipticity at wavelength *λ* (in degrees), *d* is the path length (in cm) and *c* is the protein concentration (in g/ml).

CD-measurements were made on a Jasco J-720 spectropolarimeter (Japan Spectroscopic Co., Ltd., Tokyo, Japan), calibrated according to [85]. Each CD-spectrum measurement represents the average of three scans obtained by collecting data at 1 nm intervals and was baseline-corrected by the subtraction of a spectrum obtained with the buffer only. The measurements were performed in 50 mM sodium phosphate, pH=5.2, at 24 °C in 0.1 cm quartz cuvettes. The thermal stability was determined by following the changes in the CD-signal at 222 nm while heating a protein sample from 20 to 95 °C in 50 mM sodium phosphate pH=5.2. A 0.1 cm quartz cell was used.

2.9 Cleavage of caenopore-5 by factor Xa

Caenopore-5 fusion protein (0.4 mg/ml) was incubated with factor Xa (Roche, Mannheim, Germany) at a final concentration of 10 µg/ml in 50 mM Tris-HCl, 100 mM NaCl and 1 mM $CaCl_2$ at pH=8 for 24h, at 24 °C. The correct mass of the cleavage product was confirmed by MALDI-TOF mass spectrometry (MALDI-TOF/TOF mass spectrometer 4700, Applied Biosystems, Framingham, USA).

2.10 Effect of different phospholipids on caenopore-5 conformation

The large unilamellar vesicles (LUV) were kindly provided by Matthias Michalek. They were prepared according to Ruysschaert [86]. The following lipids were used: L-α-phosphatidylcholine from egg yolk (Fluka, Taufkirchen, Germany), L-α-phosphatidyl-DL-glycerol from egg yolk (Fluka, Taufkirchen, Germany), asolectin from soybean (Fluka, Taufkirchen, Germany), a mixture of L-α-phosphatidylcholine and L-α-phosphatidyl-DL-glycerol (w/w ratio of 3:1) and a mixture of L-α-phosphatidylcholine and 3-sn-phosphatidylethanolamine from bovine brain (Fluka, Taufkirchen, Germany) (w/w ratio of 3:1). The lipids were dissolved in 1.5 ml chloroform, which was later removed by rotary evaporation under vacuum. The resulting lipid films were resuspended in 270 µl of 10 mM HEPES and 1 mM EDTA at pH=7.4 and subjected to four cycles of freezing and thawing. Next, 280 µl of 10 mM HEPES and 1 mM EDTA at pH=7.4 were added to the suspension, which was sonicated on ice at 40 % power. The liposome preparation (0.5 ml) was loaded onto a NAP-5 column (Amersham Biosciences, Freiburg, Germany), equilibrated with 10 mM HEPES, 150 mM NaCl and 1 mM EDTA at pH=7.4. For the elution, 0.75 ml of equlibration buffer were applied onto the column, but the first 0.3 ml of eluted fraction was discarded and the proper size liposome fraction was collected in the remaining eluate. For the

liposome-binding assay the liposome suspension was diluted at a ratio of 1:50 in 50 mM sodium phosphate, pH=5.2.

The interaction of caenopore-5 with different liposomes was detected by means of CD-spectroscopy. The far-UV spectra were recorded at 24 °C before and after mixing the protein solution with corresponding liposomes in a tandem quartz cuvette. The protein solution mixed with liposomes was incubated for 30 min before CD-measurement. Each CD-spectrum measurement represents the average of three scans obtained by collecting data at 1 nm intervals. The measurements were performed in 50 mM sodium phosphate, pH=5.2.

2.11 Nuclear magnetic resonance spectroscopy

2.11.1 Sample preparation for NMR measurements

The NMR experiments were performed with a solution of 1 mM – 2 mM of ^{15}N-enriched or ^{15}N/ ^{13}C-enriched caenopore-5 fusion protein in a buffer containing 50 mM sodium phosphate, 5 mM EDTA and 93 % H_2O/ 7 % D_2O at pH=5.2. In the case of the^{13}C-edited-NOESY experiment, the protein sample was in 50 mM sodium phosphate, 5 mM EDTA at pH=5.2 in D_2O

2.11.2 Two-dimensional NMR experiment

A two-dimensional-heteronuclear single quantum coherence (2D-HSQC) spectrum was recorded. NMR data were acquired at 25 °C on a Varian Inova Unity 500 MHz spectrometer (Varian, Palo Alto, USA). The spectrometer was equipped with a z-axis pulsed- field gradient $^{1}H/^{15}N/^{13}C$ probe optimised for ^{1}H detection. The experiment was recorded with the INEPT transfer delay set to 2.25 ms as a compromise between the fast relaxing amide protons and the optimal transfer time for the $^{1}J_{NH}$ couplings of 90 Hz. The spectral width was 8000 Hz and 1750 Hz for the ^{1}H and ^{15}N dimension, respectively. Data were

acquired for 312 increments of FID, each with 2048 total complex data points. Data sets were processed using the program NMRPipe [87].

2.11.3 Three-dimensional heteronuclear NMR experiments

The following 3D-heteronucear experiments were performed: ^{15}N-edited-TOCSY, ^{15}N-edited-NOESY, ^{13}C-edited-TOCSY, ^{13}C-edited-NOESY, HNCA, H(CCO)NH, C(CO)NH. The 3D-NMR spectra were utilised for the sequence-specific assignment. In addition, the distance constraints for structure calculations were obtained from ^{15}N-edited-NOESY and ^{13}C-edited-NOESY spectra. The experimental setup for all 3D-heteronuclear spectra is shown in Table 1. The ^{15}N-edited-NOESY and ^{13}C-edited-NOESY were recorded with the following mixing times: 100 ms and 120 ms, respectively. The data were further processed to obtain 3D-spectra by using the NMRPipe software [87].

2.12 Sequence-specific assignment of NMR signals

All 3D-spectra except for ^{15}N-edited-NOESY and ^{13}C-edited-NOESY were used for assigning the $^{1}H_N$, ^{1}H, ^{15}N and ^{13}C chemical shifts to a particular residue in the amino-acid sequence. The NMRView software was used for the analysis of the spectra [88].

Table 1: *Three-dimensional heteronuclear experiments acquired for caenopore–5 fusion protein, where t and F correspond to the number of complex points in the time domain and in the frequency domain of the spectra, respectively. NMR data were recorded on a Varian Inova Unity 600 MHz spectrometer (Varian, Palo Alto, USA) and a Bruker DRX 600 MHz spectrometer (Bruker GmbH, Rheinstetten, Germany).*

Experiment	Number of complex points			Spectral width (Hz)			Spectrometer
	t_1	t_2	t_3	F_1	F_2	F_3	
HNCA [89]	2048 (^{1}H)	90 (^{15}N)	120 (^{13}C)	8389	1923	4166	Bruker DRX 600
H (CCO)NH [90]	2048 (^{1}H)	42 (^{15}N)	98 (^{1}H)	8389	1923	8620	Bruker DRX 600
C(CO)NH [91]	2048 (^{1}H)	42 (^{15}N)	142 (^{13}C)	8389	1923	10638	Bruker DRX 600
^{15}N-TOCSY [92]	1024 (^{1}H)	272 (^{1}H)	64 (^{15}N)	4199	7400	1920	Varian INOVA 600
^{15}N-NOESY [92]	1024 (^{1}H)	272 (^{1}H)	64 (^{15}N)	4199	7400	1920	Varian INOVA 600
^{13}C-TOCSY [93]	1280 (^{1}H)	196 (^{1}H)	168 (^{13}C)	4194	4194	11312	Bruker DRX 600
^{13}C-NOESY [92]	2048 (^{1}H)	114 (^{13}C)	214 (^{1}H)	7507	10718	6402	Bruker DRX 600

2.13 Calculation of the caenopore–5 structure

The ^{1}H-^{1}H distance constraints derived from ^{15}N-edited-NOESY and the ^{13}C-edited-NOESY spectra were used for structure calculations performed by the DYANA software [94]. The DYANA program utilises a molecular dynamics approach in torsion angle space enhanced by simulated annealing to calculate structures consistent with the experimentally derived distances from the NMR experiments. The distance constraints were calculated from the intensity of ^{1}H resonance signals *V* [73, 95] which is related to the distance *r* between two protons by:

$$V \sim r^{-6}$$

Distances were sorted into the following ranges: 1.8 – 2.8, 2.8 – 3.4, 3.4 – 5.5 Å. Furthermore, distances between the 6 cysteines generating 3 disulphide bridges were introduced into the calculations: SG - SG (2.026 – 2.153 Å), SG - CB (3.031 – 3.133 Å) and SG - CA (2.974 – 4.494 Å). Hydrogen bond constraints were set between 1.8 – 2.4 Å and 2.8 - 3.4 Å for H-O and N-O, respectively. The structure calculation was repeated 400 times to obtain an ensemble of structures. For the final presentation, 20 structures with the lowest target functions were chosen.

2.14 Graphical representation

The programs Ribbons and GRASP were used for molecular visualisation and surface area calculation [96], [97].

3 Results

3.1 Protein expression and purification

3.1.1 Expression of caenopore-5

The expression vector pIVEX2.4a containing the caenopore-5 gene fused with an N-terminal His-tag under the control of the T7 polymerase promoter was used to transform the competent *E.coli* strain Bl21 (DE3) pABlac. The recombinant caenopore-5 fusion protein was expressed in minimal medium overnight after 1 mM IPTG had been added, which induced T7 RNA polymerase production in bacteria. Expression was analysed by SDS-PAGE. A band slightly below a molecular mass of 14.4 kDa, which corresponds to the expected molecular mass of the recombinant fusion protein which is 10.98 kDa, appeared after induction by IPTG (see Figure 6).

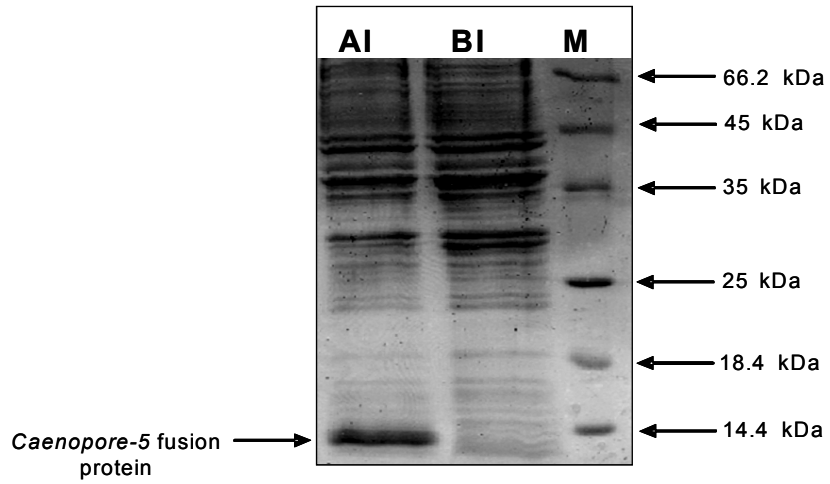


Figure 6: *The expression of caenopore-5 fusion protein in E.coli Bl21 (DE3) pABlac strain is visualised in SDS-PAGE gel: M: molecular weight marker; BI: bacteria lysate before induction with IPTG; AI: bacteria lysate after induction with IPTG.*

3.1.2 Purification of caenopore-5 from supernatant

After cell lysis, the recombinant caenopore-5 fusion protein was found in the supernatant as well as in inclusion bodies. The purification of caenopore-5 obtained from both fractions was achieved by a combination of affinity and size-exclusion chromatography, which was monitored by SDS-PAGE electrophoresis.

The supernatant containing soluble caenopore-5 was directly loaded onto a Ni^{2+}-NTA agarose column for purification, taking advantage of the presence of an N-terminal $(His)_6$ tag. The protein was efficiently bound to Ni^{2+}-NTA agarose beads, which is indicated by Figure 7. Even though the Ni^{2+}-NTA agarose beads with bound caenopore-5 were washed intensively with an appropriate buffer to remove all unspecifically bound material, the protein present in the eluted fractions was not sufficiently pure (see Figure 7). Therefore, the eluted fractions were applied on a size-exclusion chromatography column for further purification (see Section 3.1.4).

3.1.3 Refolding and purification of caenopore-5 from inclusion bodies

The insoluble and denatured protein obtained as inclusion bodies was subjected to refolding *in vitro* by overnight dialysis, after solubilisation in 6 M Gdn·HCl. Under these refolding conditions, almost 80 % of the recombinant protein remained soluble in the used buffer system and the fold of the soluble protein was verified by CD-spectroscopy as described in Section 2.8. Furthermore, the SDS-PAGE analysis of the refolded protein showed that it contained many impurities. It was therefore applied onto a Ni^{2+}-NTA agarose column for purification (see Figure 8). The examination of the eluted fractions by SDS-PAGE revealed that they contained purified protein with only a small amount of contamination in comparison to the protein from the supernatant. However, these fractions were also applied onto a Hi-Load Superdex 75 column in order to improve the purity of the sample.

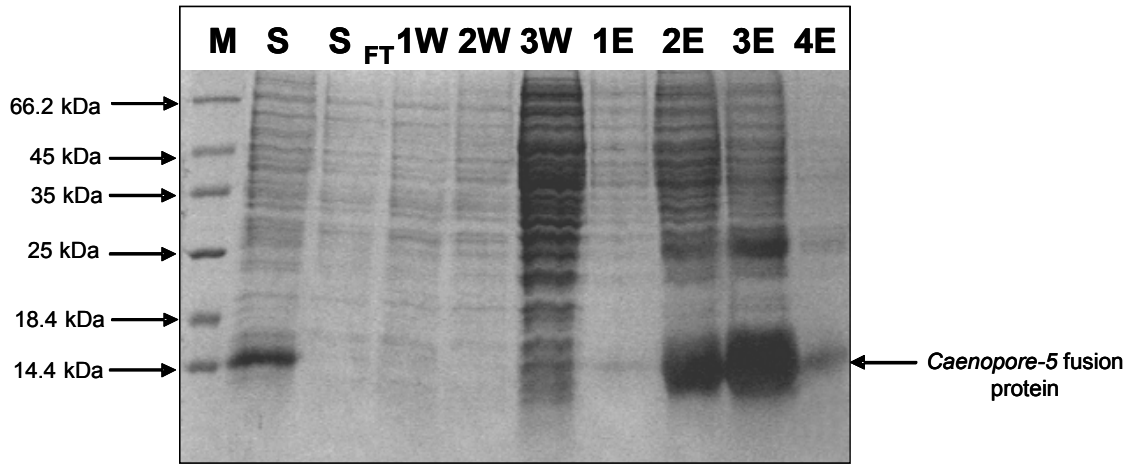


Figure 7: *SDS-PAGE analysis of the purification of caenopore-5 fusion protein from supernatant by affinity chromatography (Ni^{2+}-NTA agarose): M: molecular weight marker; S: supernatant fraction with caenopore-5 fusion protein before loading onto Ni^{2+}-NTA agarose beads; S_{FT}: supernatant flow through; W: washing fractions of Ni^{2+}-NTA agarose beads (1 ml); E: elution fractions (1 ml).*

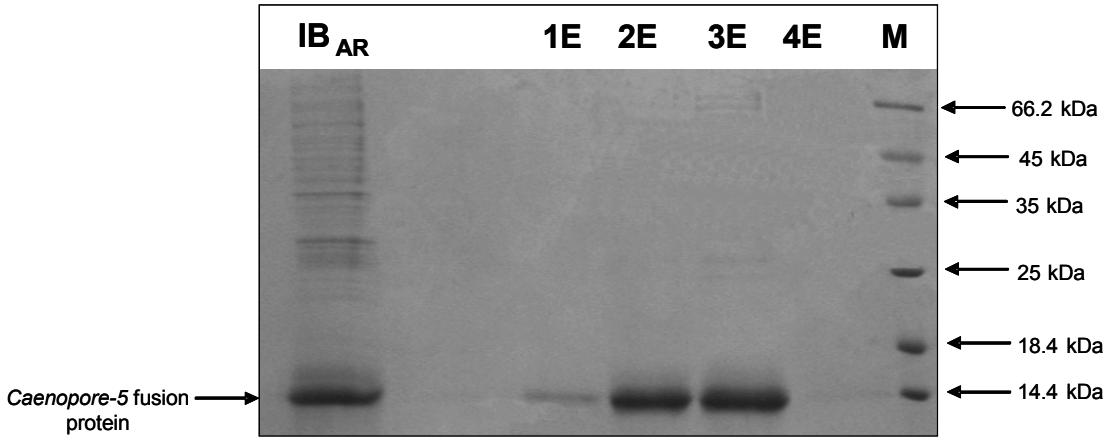


Figure 8: *SDS-PAGE analysis of caenopore-5 fusion protein purity derived from inclusion bodies, which was refolded by dialysis and purified by Ni^{2+}-NTA agarose: IB_{AR}: soluble inclusion bodies fraction after refolding; E: elution fractions (1 ml) from Ni^{2+}-NTA agarose; M: molecular weight marker.*

3.1.4 Size-exclusion chromatography

A Hi-Load Superdex 75 column was calibrated with a mixture of proteins with a known molecular weight ranging from 6.5–75 kDa. The calibration was carried out in 50 mM sodium phosphate and 5 mM EDTA at pH=5.2. The logarithm of the molecular weight (MW) of the standard marker proteins (y axis) was plotted versus the retention volume (x axis) to give a calibration curve described by the following equation:

$$y=-0.0286x + 6.4481$$

The calibration curve is represented in Figure *9*.

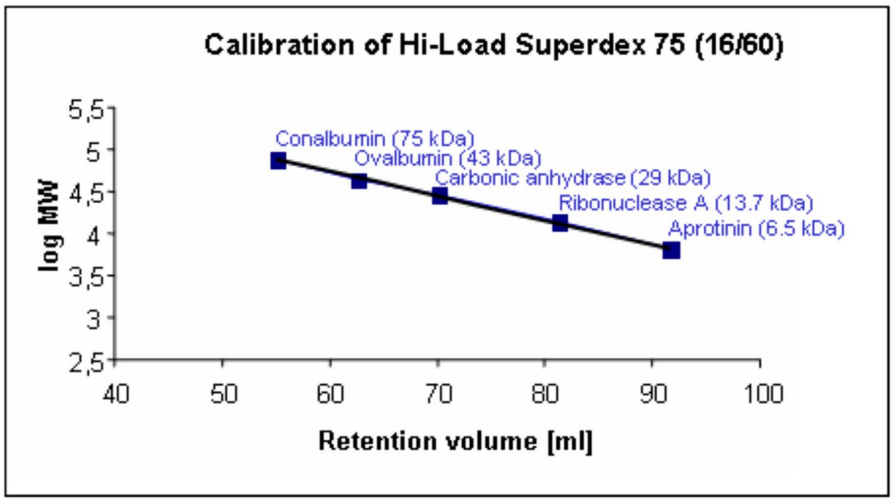


Figure 9: *The calibration curve depicting the logarithm of MW (molecular weight of standard markers) versus the retention volume (blue squares). The calibration was performed in 50 mM sodium phosphate and 5 mM EDTA at pH=5.2.*

On the basis of the calibration equation, the elution volume of 84 ml for caenopore-5 fusion protein was calculated using a molecular weight of 10.98 kDa.

Fractions derived from the affinity chromatography (see Sections 3.1.2 and 3.1.3) containing caenopore-5 fusion protein were applied onto a size-exclusion column, equlibrated with 50 mM sodium phosphate and 5 mM EDTA at pH=5.2. The elution profile is represented in Figure 10. Caenopore-5 fusion protein was

eluted at a retention volume equal to 81.25 ml, which corresponds to a monomeric state of the protein, but does not conform to the value calculated on the basis of the calibration equation. This discrepancy in the elution volume of caenopore-5 might be explained by the presence of the N-terminal fusion $(His)_6$-tag, which as a flexible extension may increase the size of molecule, compared to the globular shape of standard proteins used for calibration. Additionally, the SDS-PAGE examination under reducing and non-reducing conditions of the eluted protein indicates the purification of caenopore-5 fusion protein to homogeneity (see Figure 10). The higher mobility of the protein in non-reducing conditions is an indicator for the more compact molecule conformation due to the intact disulphide bridges.

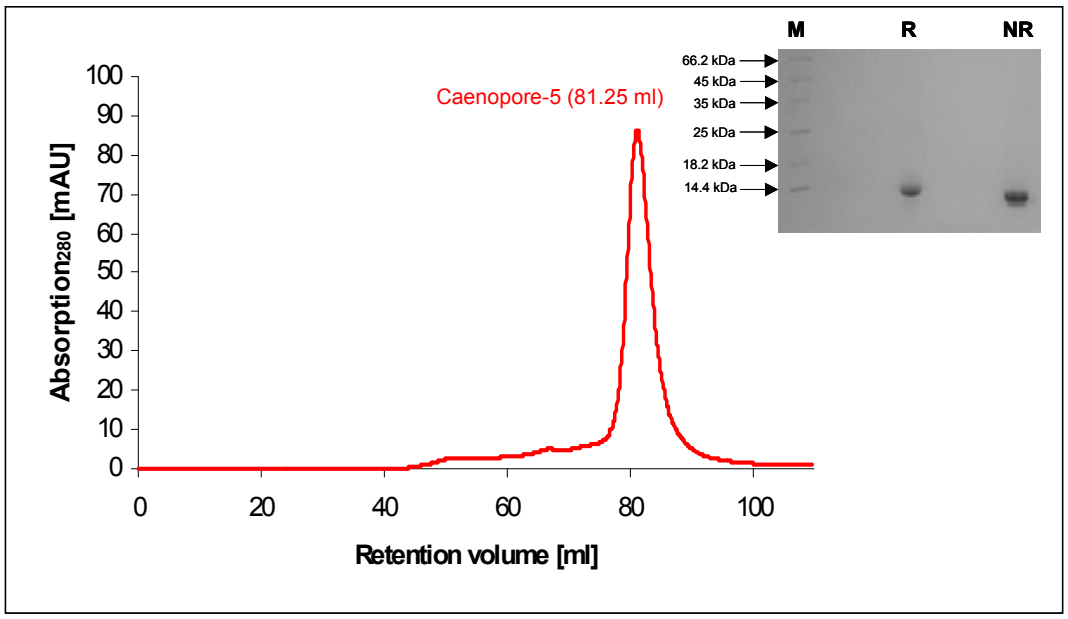


Figure 10: *Elution profile of caenopore-5 fusion protein in 50 mM sodium phosphate and 5 mM EDTA at pH=5.2 on a Hi-Load Superdex 75 (16/60). Caenopore-5 fusion protein was eluted at a retention volume of 81.25 ml. The inserted image represents the SDS-PAGE purity analysis of the caenopore-5 fusion protein after the final purification step by size-exclusion chromatography: M: molecular weight marker; R: caenopore-5 fusion protein under reducing conditions; NR: caenopore-5 fusion protein under non-reducing conditions.*

3.2 Circular dichroism spectroscopy

The secondary structure of the purified caenopore-5 fusion protein obtained from supernatant and inclusion bodies was examined by far-UV CD-spectroscopy.

Caenopore-5 fusion protein showed the CD-spectrum typical for alpha-helical fold with two characteristic minima at 208 nm and 222 nm, and a maximum at 195 nm (see Figure 11). The spectrum is characteristic for alpha-helical proteins and is also observed for other members of the SAPLIP family. Furthermore, the thermal stability of the protein was tested by measuring the intensity of the CD-signal at 222 nm while increasing the temperature. Caenopore-5 fusion protein appeared to be resistant to heating up to 95 °C as only very subtle changes of the signal at 222 nm were observed (data not shown). This result was confirmed by the far-UV spectrum of the protein after being subjected to the thermal stress (see Figure 12). Caenopore-5 fusion protein did not seem to undergo any modification, proving its extreme stability.

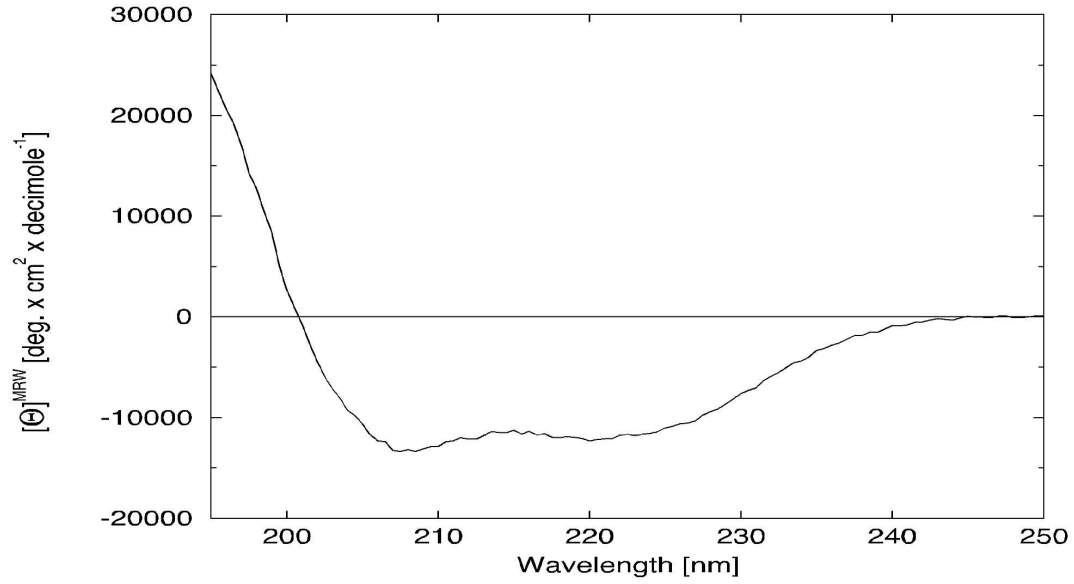


***Figure 11**: Far-UV CD-spectrum of caenopore-5 fusion protein, recorded in 50 mM sodium phosphate at pH=5.2, at 24 °C.*

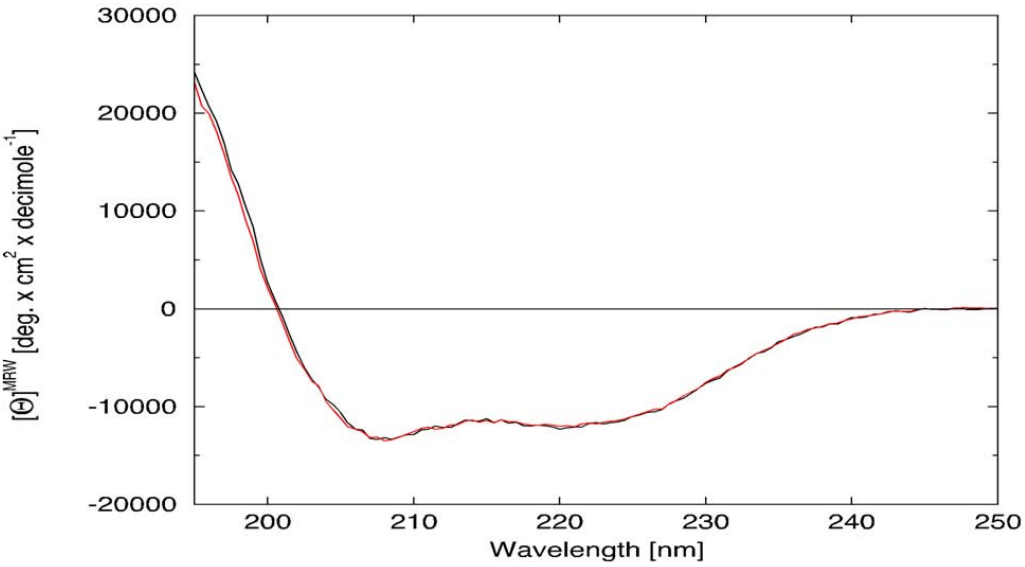


Figure 12: *Far-UV CD-spectra of caenopore-5 fusion protein, recorded in 50 mM sodium phosphate at pH=5.2, at* 24 °C*. Black line: CD-spectrum before increasing temperature to 95 °C; red line: CD-spectrum after heating up to 95 °C.*

3.3 Sequential assignment

The sequential assignment is the prerequisite for solving the structure of a protein by using the NMR spectroscopy. It deals with the assignment of each signal (resonance) present in the spectra to the particular nucleus in the amino-acid sequence. The following heteronuclear 3D spectra were used for the sequence-specific assignment: HNCA, C(CO)NH, H(CCO)NH, and ^{15}N-edited TOCSY. The resonances are resolved according to the three dimensions corresponding to ^{15}N, ^{1}H directly attached to ^{15}N (^{1}HN, amide proton) and ^{13}C or ^{1}H atom chemical shifts. The common feature of all of these spectra is the correlation of the ^{1}HN with the ^{13}C or ^{1}H atoms of its own residue or of the preceding one, depending on the type of the spectrum. Furthermore, all of these experiments employ a magnetisation transfer via chemical bonds utilising a one, two or three-bond scalar couplings. The initial step in the sequential assignment procedure was the attribution of each backbone ^{1}HN signal to a particular amino-acid residue in the caenopore-5 sequence, followed by the assignment of the ^{1}H and ^{13}C resonances of the side chains.

3.3.1 Sequential assignment of amino-acid residues

An HNCA spectrum correlates the ^{1}HN with the ^{13}CA (carbon alpha) of its own residue (i) and of the preceding one (i - 1), which enables the generation of sequential connectivity along a protein's sequence (see Figure 13).

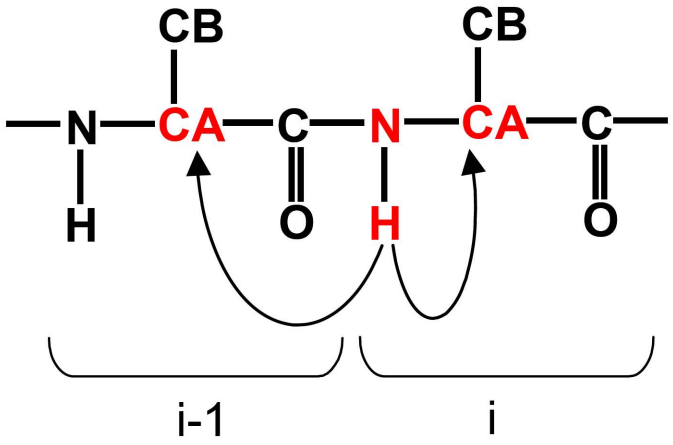


***Figure 13**: Schematic representation of the sequential connectivity generated by the 3D-HNCA experiment. The arrows depict the correlation of a ^{1}HN with the ^{13}CA atom of its own residue (i) and of the preceding one (i-1).*

This connectivity can be interrupted by proline residues, which lack a ^{1}HN proton. At each ^{1}HN/^{15}N resonance pair, two ^{13}CA resonances are usually present in the ^{13}C frequency dimension and in most cases their different intensities helped to discriminate between the intraresidual and the interresidual signals (see Figure 14). By analysing the HNCA spectrum, stretches of amino-acid residues that are sequentially connected were identified. As a next step, they had to be assign to corresponding fragments within the caenopore-5 sequence.

To identify the type of amino-acid residue that gave rise to a signal in the HNCA experiment, a C(CO)NH spectrum was utilised. The C(CO)NH spectrum correlates a ^{1}HN/^{15}N pair of a given residue (i) with the ^{13}C atoms of its preceding residue (i - 1), which thereby creates the connectivity between neighbouring amino acids (see Figure 15).

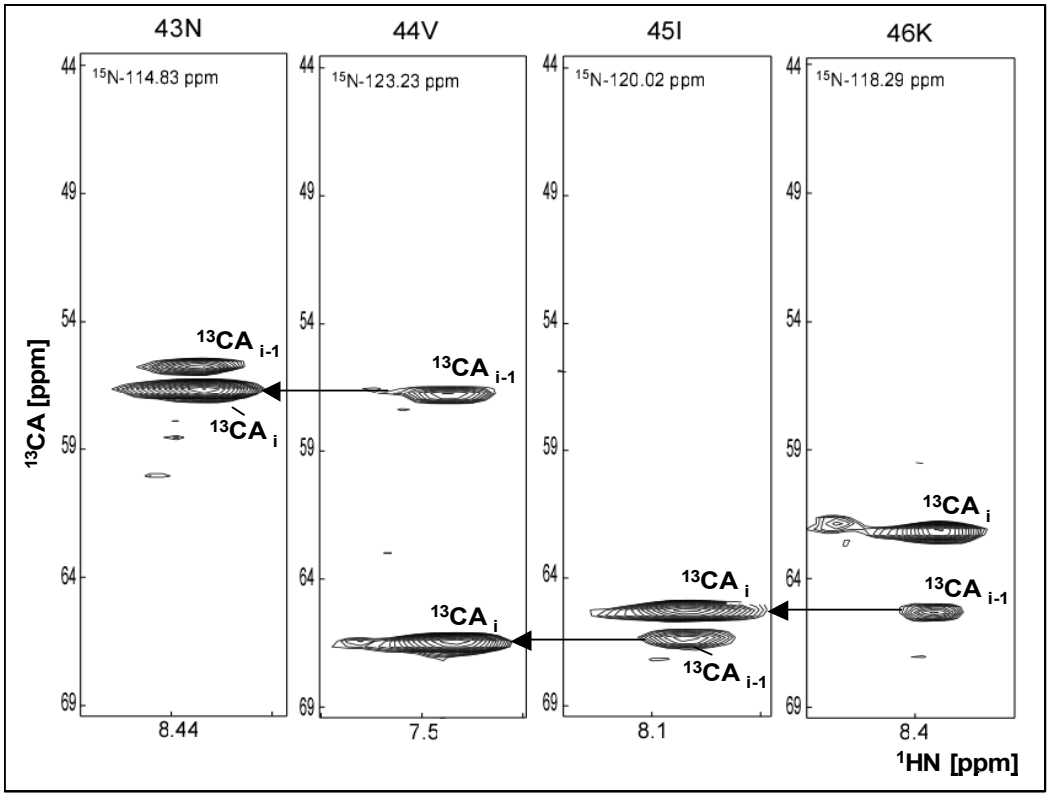


Figure 14: *Sequential connectivity generated by the 3D-HNCA experiment depicted for an assigned fragment of caenopore-5 fusion protein. Each 2D section derived from different ^{15}N planes of the 3D-HNCA spectrum represents a single residue. They show the ^{13}CA resonances of a given residue (i) and its preceding residue (i-1) at different chemical shift values. The arrows indicate the connectivity between sequential amino acids provided by the ^{13}CA resonance signal of the preceding residue.*

The number of ^{13}C resonances and their chemical shifts, which are characteristic for a particular ^{13}C atom (^{13}C alpha, beta, gamma or delta) (see Figure 16), lead to the identification of the amino-acid residue. It is shown in Figure 17 that each residue possesses a pattern of specific ^{13}C resonances that is distinctive for the side chain of the preceding amino acid. Additionally, the comparison of HNCA and C(CO)NH signals of the same residue contributed to the distinction between intraresidual and interresidual ^{13}CA resonances in the HNCA spectrum (see Figure 14 and Figure 17).

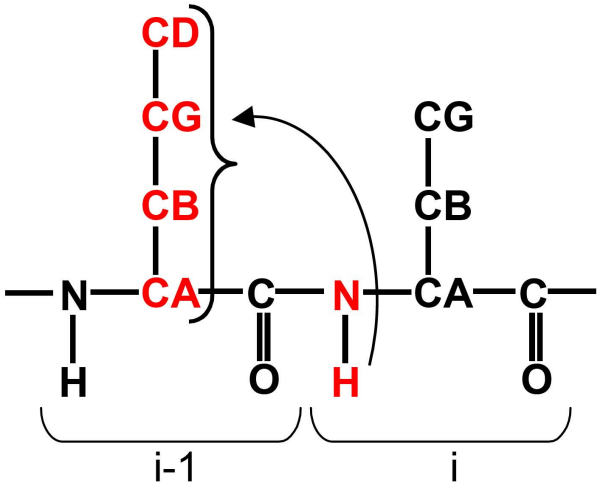


Figure 15: *Schematic representation of the sequential connectivity detected in a 3D-C(CO)NH spectrum. The arrow depicts the correlation of the ^{1}HN of a given residue (i) with ^{13}C atoms of a preceding one (i - 1).*

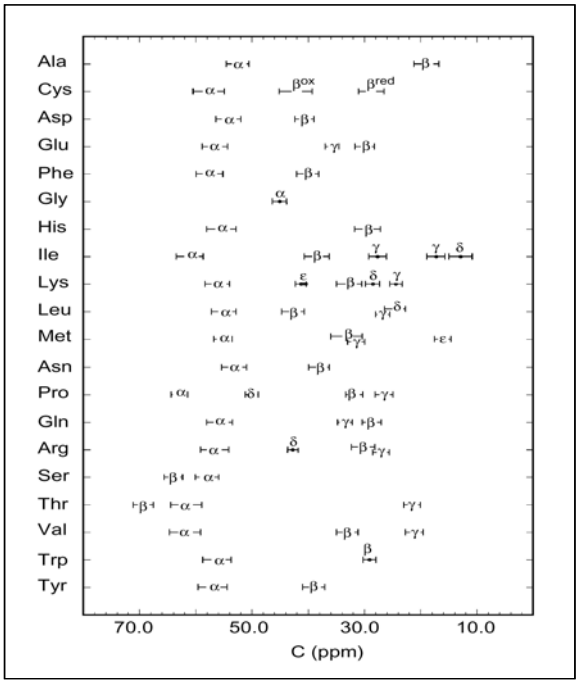


Figure 16: *Representation of the ^{13}C chemical shifts [67]. The mean values for aliphatic carbons (α, β, γ or δ) in each amino-acid type were obtained from the data base of 13 proteins described by Kraulis [98]. The bars represent the standard deviation from the mean value.*

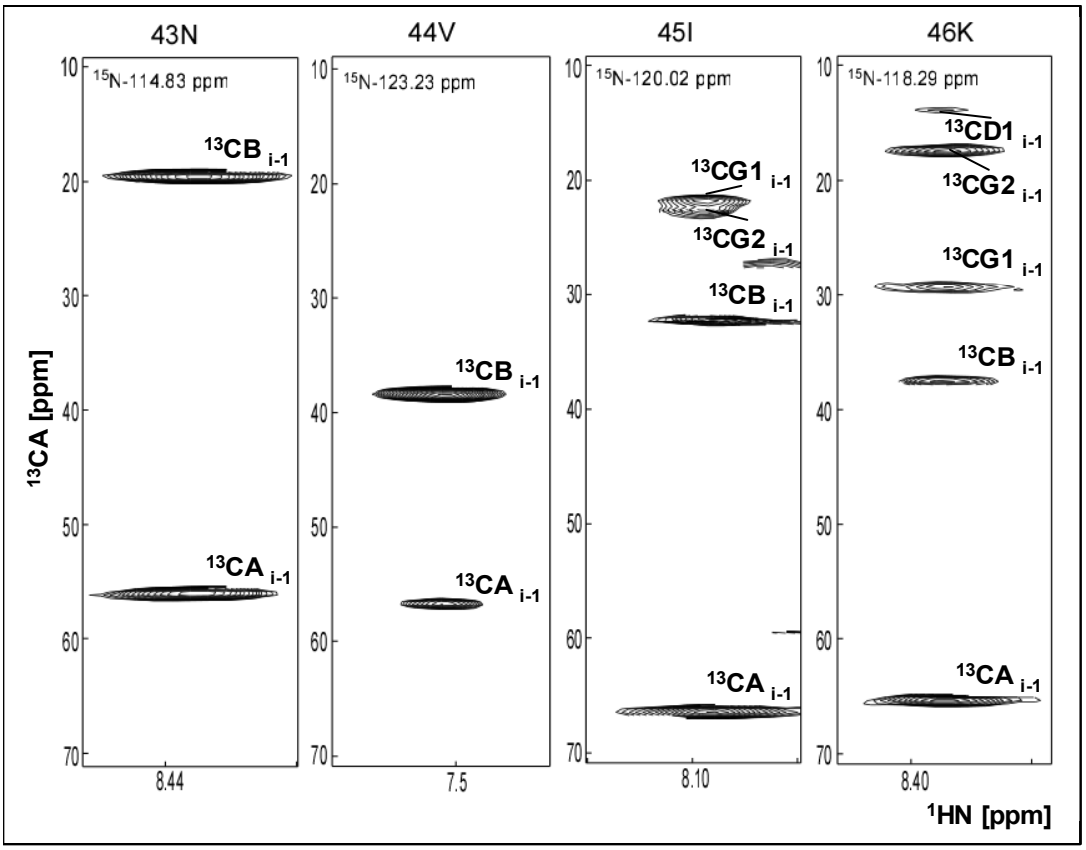


Figure 17: *The 2D sections derived from different ^{15}N planes of the 3D-C(CO)NH spectrum, which represent single residues from assigned fragment of caenopore-5 fusion protein (the same residues as in figure 14). They show ^{13}C resonances of preceding amino-acid residues (i-1) at different chemical shift values.*

Regarding the ambiguity of ^{13}CA resonances in a HNCA spectrum, two other spectra: a 3D-^{15}N-edited TOCSY and a 3D-H(CCO)NH were used to support the sequential assignment. Both the ^{15}N-edited TOCSY and the H(CCO)NH spectrum correlate ^{1}HN with ^{1}H atoms. The former, however, correlates only ^{1}H atoms of its own residue (i), whereas the latter correlates ^{1}H resonances of the preceding amino acid (i - 1) (see Figure 18). The combination of these two spectra accomplished the sequential connectivity (see Figure 19). Furthermore, the number of ^{1}H resonances and their chemical shifts were also used for the identification of the amino-acid type.

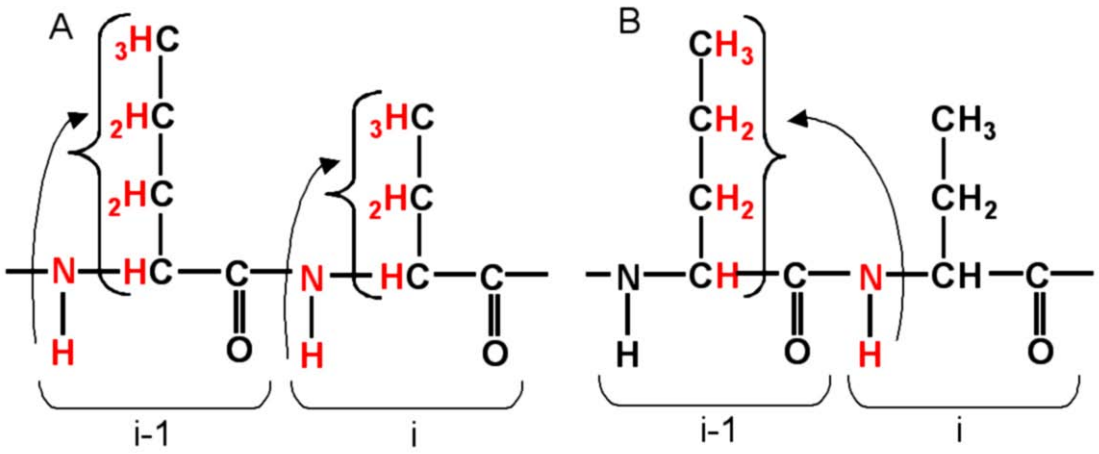


Figure 18: *Schematic representation of the sequential connectivity generated by combining a 3D-^{15}N-edited TOCSY* ***(A)*** *and a 3D-H(CCO)NH* ***(B)****. The arrows indicate the correlation of ^{1}HN with ^{1}H atoms of its own residue (i) or of a preceding one (i-1) in a 3D-^{15}N-edited TOCSY and a 3D-H(CCO)NH, respectively.*

The result of the ^{1}HN-resonance assignment to single amino-acid residues in the caenopore-5 sequence (including the N-terminal tag) is depicted in the 2D-^{1}H-^{15}N HSQC spectrum, which represents the ^{15}N/^{1}HN projection of all above-mentioned spectra (see Figure 20). In the 2D-^{1}H-^{15}N HSQC spectrum each signal represents a ^{1}H nuclei attached covalently to a ^{15}N atom. Therefore, all amino-acids in the protein sequence contribute to the signals in the 2D-^{1}H-^{15}N HSQC spectrum, except for the proline. The number of expected backbone ^{1}HN signals in the HSQC spectrum is 94 (17 derived from N-terminal tag fused with the caenopore-5 sequence and 77 present in the native caenopore-5 sequence). Apart from ^{1}HN of Phe 62, all expected resonances from the native caenopore-5 sequence (residues: 19 - 99) were assigned. Surprisingly, 20 residues had double backbone amide proton chemical shifts (see Figure 20), indicating that one ^{1}H/^{15}N pair exists in two distinct environments, e.g. the presence of two different conformers. These double peaks are present for the following residue ranges: 19 - 27, 60 - 66 and 90 - 98. Furthermore, the intensities of these two sets of signals are similar, which is indicative for the same ratio of these two conformers. As opposed to caenopore-5 fusion protein sequence, most of the amide protons representing the N-terminal tag (residues: 1 - 18) were not observed in the spectra. Only 7 of 17 expected

backbone resonances from this region were successfully assigned. In total, 104 backbone signals were assigned achieving nearly complete assignment of signals which appeared in the 2D-^{1}H-^{15}N HSQC spectrum. In addition, the resonances representing amide groups from the side chains of asparagine and glutamine residues were present in the spectrum. The side-chain amide protons of Gln 25 and Asn 96 were also doubled.

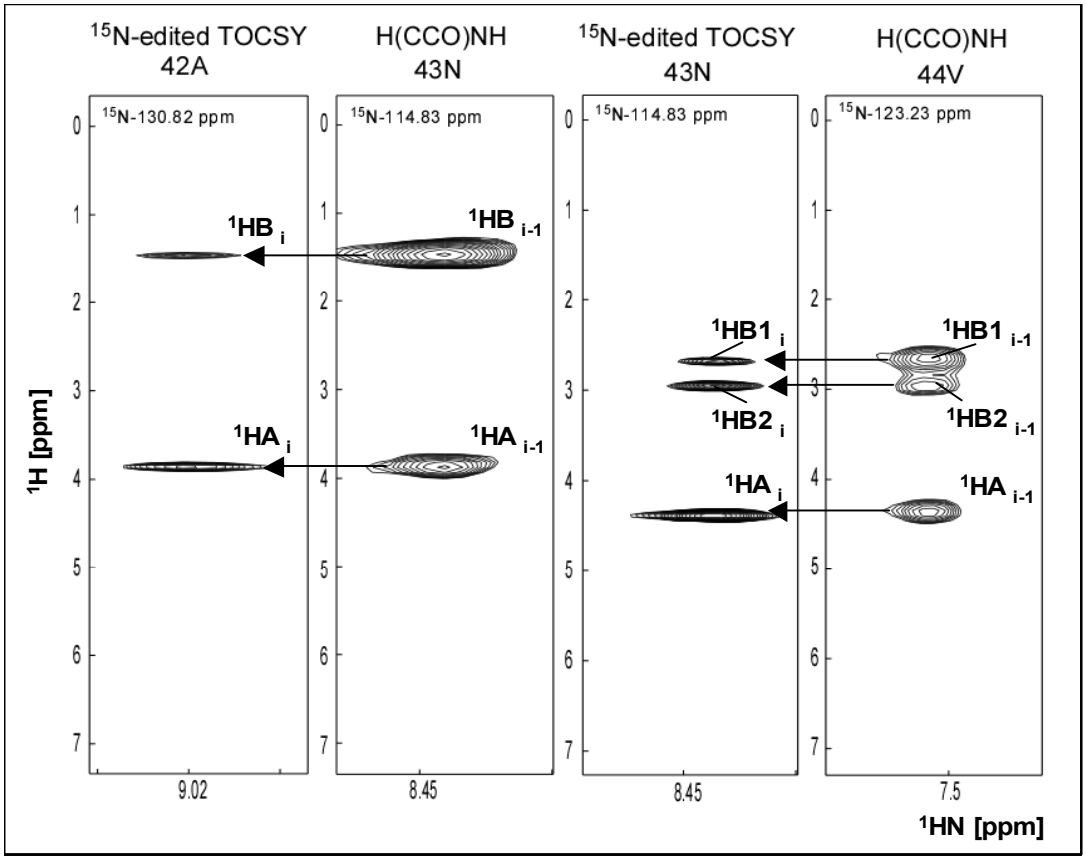


Figure 19: *2D sections derived from different ^{15}N planes of the 3D-^{15}N-edited TOCSY and 3D-H(CCO)NH spectra, which represent single residues from an already assigned fragment of caenopore-5 fusion protein. They show ^{1}H resonances of a given amino acid (^{15}N-edited TOCSY) or of a preceding one (H(CCO)NH) at different chemical shift values. The sequential connectivity established by the combination of both spectra is indicated by arrows.*

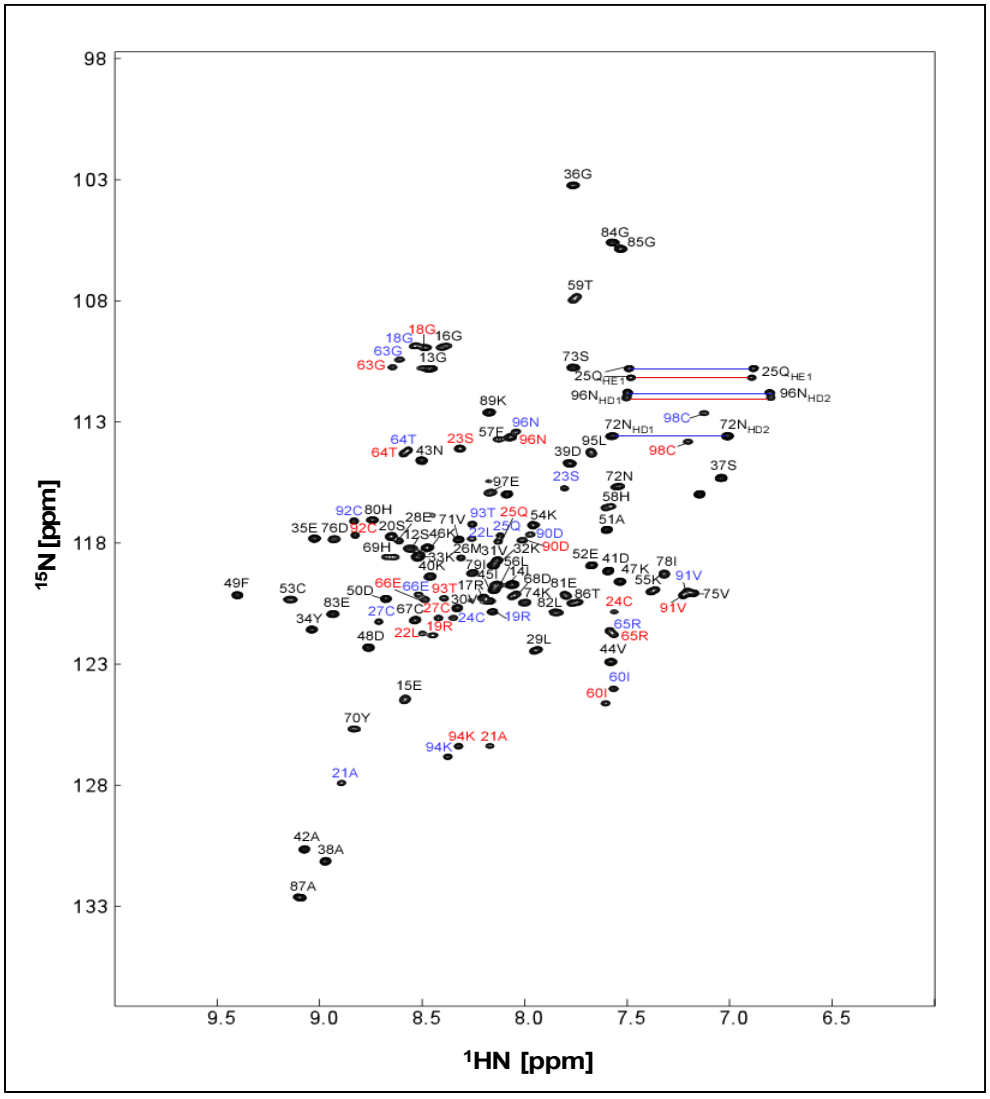


Figure 20: *The 2D-^{1}H-^{15}N HSQC spectrum of caenopore-5 fused with the N-terminal tag (1-99 residues). Assigned resonances are labelled with the corresponding amino-acid residue. Residue names coloured in blue and red belong to conformer I and conformer II, respectively. The blue and red lines connect amide protons assigned to the side chains of glutamine and asparagines amino acids of conformer I and conformer II, respectively.*

3.3.2 Assignment of amino-acid side chain resonances

The assignment of ^{1}H or ^{13}C atoms of the amino-acid side chains followed the allocation of the ^{1}HN resonances to particular residues in the caenopore-5 fusion protein sequence. The above-mentioned spectra, the ^{15}N-edited TOCSY, the H(CCO)NH and the C(CO)NH provided information about the chemical shift values of the ^{1}H or ^{13}C atoms, which was utilised to interpret the distinct type of spectrum, the 3D-^{13}C-edited TOCSY. Instead of the ^{15}N chemical shifts, the resonances in the 3D-^{13}C-edited TOCSY spectrum are resolved along the ^{13}C frequency axis. This spectrum correlates ^{1}H atoms which are directly attached to particular ^{13}C (^{1}HC) to other intraresidual ^{1}H atoms, which significantly improves the side chain assignment (see Figure 21).

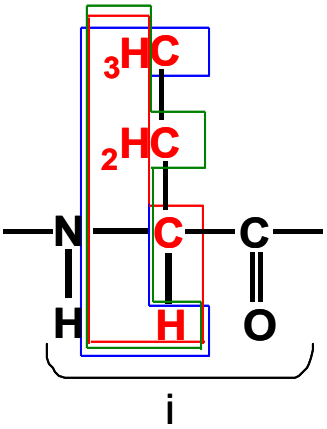


Figure 21: *Schematic representation of the intraresidual relationship in a 3D-^{13}C-edited TOCSY. The correlation of a ^{1}H atom or atoms directly attached to a particular ^{13}C type with all intraresidual ^{1}H atoms is indicated by the frame. Different colours of the frames correspond to various intraresidual ^{13}C types.*

At each frequency of a ^{1}H atom, which is directly bound to a particular ^{13}C type, the resonances of all intraresidual ^{1}H are present. In the example shown in Figure 22, the same pattern of ^{1}H resonances is observed in two different ^{13}C frequency planes corresponding to ^{13}CA and ^{13}CB of Asn 43.

In the case of the ^{13}CB frequency plane, the two rows of ^{1}H signals at distinct chemical shifts are present, indicating two ^{1}HB atoms bound directly to ^{13}CB.

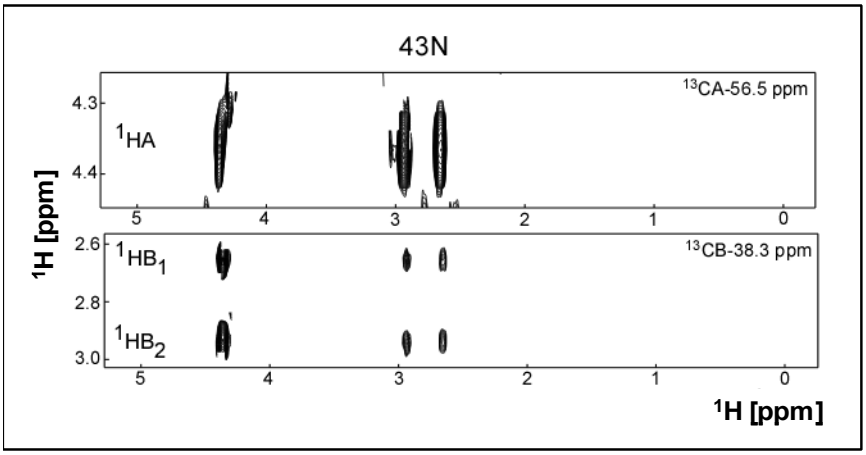


Figure 22: *2D sections derived from different ^{13}C planes of a single residue (N 43) in a 3D-^{13}C-edited TOCSY spectrum. The intraresidual ^{1}H resonances are present at the chemical shift value of a ^{1}H atom directly bound to the corresponding ^{13}C.*

3.4 Determination of the secondary and the tertiary structure of caenopore-5

The assignment of most of the resonances to particular nuclei in the caenopore-5 molecule formed the basis of the further identification of the connectivity between ^{1}H atoms which are in close proximity due to the protein's three-dimensional structure.

3D-^{15}N-edited NOESY and 3D-^{13}C-edited NOESY spectra represent the contacts between ^{1}H atoms which are less than 5 Å apart from each other as their signals are generated by through-space dipole coupling. Therefore, they depict ^{1}H resonances derived from residues which are either sequential neighbours or are far apart within the protein sequence, but are close to each other in space. Additionally, the intensity of ^{1}H resonances in the NOESY spectra is proportional to the inverse sixth power of the internuclear distance (see Section 2.13), which was used to compute ^{1}H-^{1}H distance constraints.

The CD-measurements confirmed that caenopore-5 accommodates a mostly alpha-helical conformation. This type of secondary structure is characterised by the presence of hydrogen bonds between each backbone HN and a carbonyl group of an amino acid located four residues away. This structural conformation implicates the occurrence of a specific pattern of signals in both NOESY spectra [99] .

The ^{15}N-edited NOESY spectrum correlates each ^{1}HN atom with all ^{1}H atoms present in spatial proximity (see Figure 23). The close distance between ^{1}HN and ^{1}HA atoms of the preceding third (i - 3) and fourth (i - 4) residue in alpha-helical segments contributes to characteristic resonances present in a ^{15}N-edited NOESY which are indicated by Figure 24.

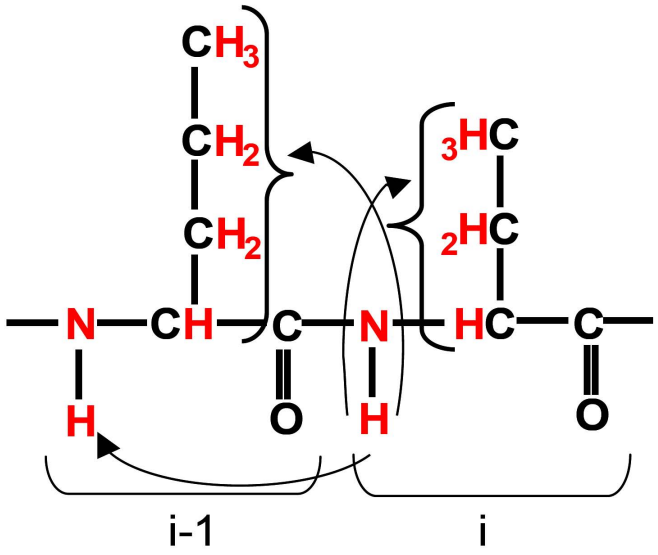


Figure 23: *Schematic representation of intraresidual and interresidual relationships in a 3D-^{15}N-edited NOESY. The arrows indicate the correlation of ^{1}HN with all ^{1}H atoms present in its close proximity.*

As a 3D-^{15}N-edited NOESY spectrum contains many ^{1}H signals for a given residue (i), the discrimination between intraresidual and interresidual atoms was accomplished by comparing these signals to the signals obtained for the same residue in a 3D-^{15}N-edited TOCSY spectrum or to the signals obtained for the following residue (i + 1) in the 3D-H(CCONH) spectrum (see Figure 24).

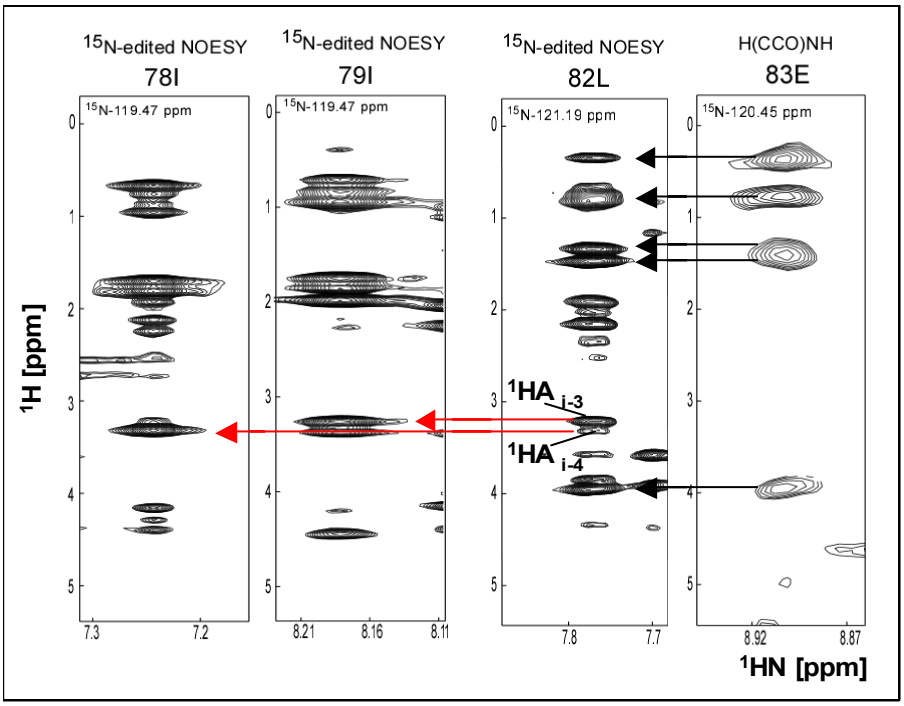


Figure 24: *2D sections derived from different ^{15}N planes of the 3D-^{15}N-edited NOESY and 3D-H(CCO)NH spectra, which represent single residues from already assigned fragments of caenopore-5 fusion protein. They show ^{1}H resonances which are in close proximity to the HN of a given amino acid (^{15}N-edited NOESY) or ^{1}H resonances of a preceding residue (H(CCO)NH). The intraresidual ^{1}H resonances in the ^{15}N-edited NOESY are indicated by black arrows. The signals characteristic for an alpha-helical segment (^{1}HN, $^{1}HA_{i-3}$), (^{1}HN, $^{1}HA_{i-4}$) are indicated by red arrows.*

The 3D-^{13}C-edited NOESY spectrum correlates a ^{1}H atom bound directly to a particular ^{13}C (^{1}HC) with other ^{1}H atoms in its close spatial proximity. Therefore, the close contact between a ^{1}HB of a given residue (i) and a ^{1}HA of an amino acid four residues apart (i - 4) within the protein sequence, which is characteristic for an alpha-helical stretch, was identified in the ^{13}C-edited NOESY spectrum (see Figure 25). A comparison with the 3D-^{13}C-edited TOCSY made it possible to distinguish between intraresidual and interresidual signals (see Figure 25).

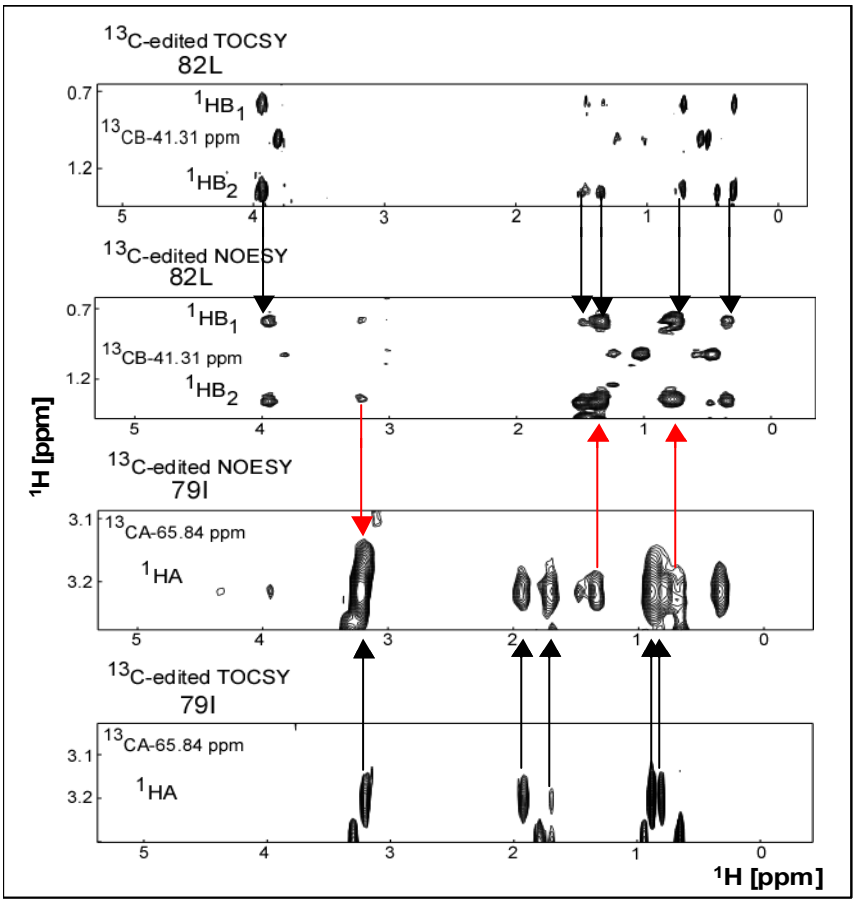


Figure 25: *2D sections derived from different ^{13}C planes of single residues in the 3D-^{13}C-edited TOCSY and 3D-^{13}C-edited NOESY spectra. The ^{1}H resonances are present at the chemical shift value of a ^{1}H atom bound directly to a corresponding ^{13}C. The intraresidual ^{1}H resonances are indicated by black arrows in ^{13}C-edited NOESY spectra. The red arrows indicate a ^{1}HA resonance of I 79 present at the chemical shifts of two ^{1}HB of L 82 and vice versa.*

To define the tertiary structure of caenopore-5, the interpretation of the 3D-^{13}C-edited NOESY spectrum predominated over the 3D-^{15}N-edited NOESY. The ^{13}C-edited NOESY spectrum provides much more detailed information about interactions between ^{1}H atoms derived from the amino-acid side chains, which could not be obtained from contacts of backbone ^{1}HN in the ^{15}N-edited NOESY. An example of such an interaction between a ^{1}HD of the long side chain of Ile 79 and ^{1}HB of Ala 42 is shown in Figure 26. Both the ^{1}HD of Ile 79 and the ^{1}HB of Ala 42 have many signals but only some of them are

intraresidual, while the other signals result from the spatial proximity of these atoms. The ^{1}HD of Ile 79 possesses a resonance with the chemical shift of ^{1}HB of Ala 42 and *vice versa,* which indicates the close distance between these two atoms.

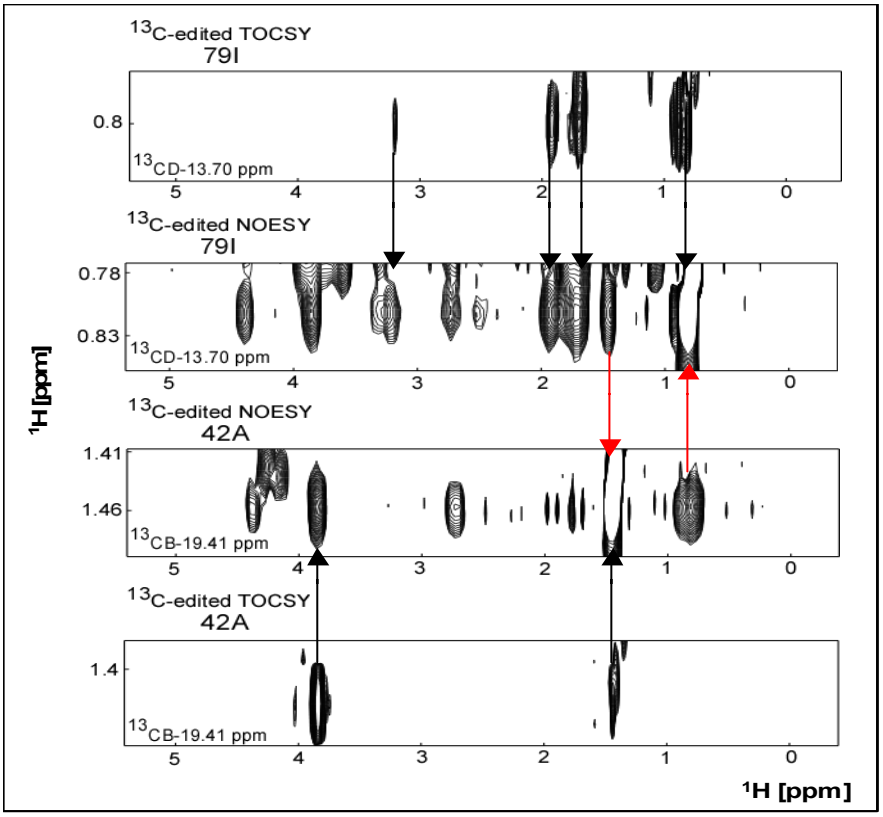


Figure 26: *2D sections derived from different ^{13}C planes of single residues in the 3D-^{13}C-edited TOCSY and the 3D-^{13}C-edited NOESY spectra. The ^{1}H resonances are at the chemical shift value of a ^{1}H atom bound directly to a corresponding ^{13}C. The intraresidual ^{1}H resonances are indicated by black arrows in the ^{13}C-edited NOESY spectrum. The red arrows indicate ^{1}HB resonance of A 42 present at the chemical shifts of ^{1}HD of I 79 and vice versa.*

3.5 Inspection of residues with double backbone amide chemical shifts

As mentioned in Section 3.3.1, the residues: 19 - 27, 60 - 66 and 90 - 98, are represented by two different backbone ^{1}HN/^{15}N resonances. To enssure the correct sequence assignment of these double resonances to the corresponding conformers of caenopore-5, a detailed analysis of sequential contacts in the ^{15}N-edited NOESY spectrum was performed. Furthermore, an examination of the ^{15}N-edited NOESY spectrum showed that these double residues have a similar pattern of assigned ^{1}H resonances with only slight differences in their signal intensities (see Figure 27). The comparison of the ^{15}N-edited NOESY with the ^{15}N-edited TOCSY spectrum for the same residue enabled to discriminate between intraresidual and interresidual ^{1}H resonances. Noteworthy was the presence of contacts between ^{1}HN and ^{1}HA of the preceding third (i - 3) and fourth (i - 4) residue particularly for these residues located in alpha-helical regions. For example, both sets of ^{1}H resonances representing Cys 27 in the ^{15}N-edited NOESY spectrum possessed signals from ^{1}HA of Cys 24 and Ser 23 (see Figure 27), indicating no perturbations in this alpha-helical fragment for these two forms of caenopore-5 fusion protein.

Such a conformational heterogeneity due to cis-trans isomerisation of a proline residue was described for other proteins [100], [101], [102]. To explain the existence of these two conformers in caenopore-*5* fusion protein, the cis-trans isomerisation of proline residues in regions with double chemical shifts for backbone amide protons were therefore further investigated. The short proton-proton distances between proline and its preceding residue were analysed in the ^{13}C-edited NOESY as they are indicators of changes in backbone conformation resulting from proline isomerisation [68].

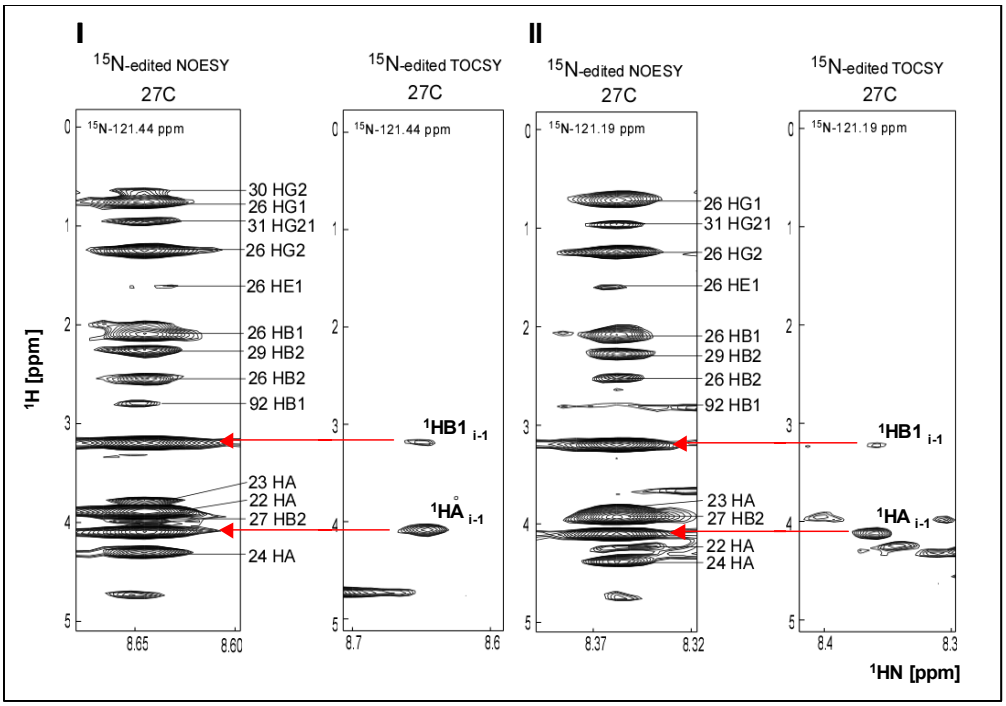


Figure 27: *2D sections derived from ^{15}N planes of the 3D-^{15}N-edited NOESY and the 3D-^{15}N-edited TOCSY spectra of the C 27 residue represented by two different backbone amide proton chemical shifts: I (8.65 ppm) and II (8.36 ppm). They show resonances of protons which are in close proximity to the HN of C 27 (^{15}N-edited NOESY) or the intraresidual ^{1}H resonances of C 27 (^{15}N-edited TOCSY). The intraresidual ^{1}H resonances in the ^{15}N-edited NOESY are indicated by red arrows (except for HB2 of C 27, which was assigned by means of ^{13}C-edited TOCSY). All assigned ^{1}H resonances in the ^{15}N-edited NOESY are labelled with amino-acid residue number and proton type.*

Furthermore, ^{1}H resonances of proline residue could only be observed in the ^{13}C-edited NOESY and the ^{13}C-edited TOCSY spectra. The experimental data provided evidence for the cis-trans isomerisation only for Pro 99 (see Figure 28). The presence of the contact between ^{1}HA of Pro 99 and ^{1}HA of Cys 98, which is characteristic for a cis peptide bond conformation, could be observed in one conformer of caenopore-5 fusion protein (conformer I). In contrast, the second form of caenopore-5 fusion protein only showed the ^{1}HA resonance belonging to Cys 98 at chemical shifts of two ^{1}HD protons of Pro 99

indicating their close interaction which is typical for trans peptide bond conformation (conformer II).

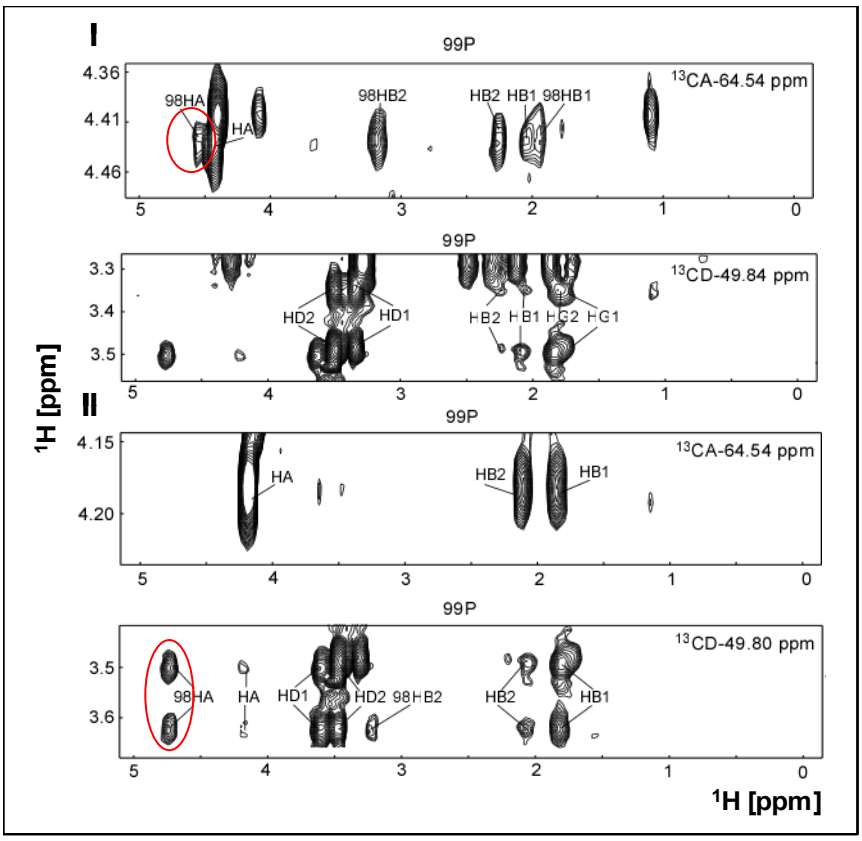


Figure 28: *2D sections derived from ^{13}CA and ^{13}CD planes of ^{13}C-edited NOESY spectrum of the P 99 belonging to two different conformers of caenopore-5 fusion protein:* ***I*** *and* ***II****. The ^{1}H resonances are at the chemical shift value of a ^{1}H atom bound directly to a corresponding ^{13}C. The intraresidual ^{1}H resonances are only labelled with proton type and the interresidual ^{1}H resonances are additionally provided with a residue number. The ^{1}HA of C 98, which is of particular interest here, is circled.*

3.6 Structure description

The structure of caenopore-5 was determined on the basis of ^{1}H-^{1}H distance constraints derived from the ^{15}N-edited-NOESY and the ^{13}C-edited-NOESY spectra. Caenopore-5 is a monomeric peptide with alpha-helical secondary structural elements (see Sections 3.1.4 and 3.2), resulting in the presence of characteristic medium-range ^{1}H-^{1}H resonances in the spectra: $d_{\alpha N}(i,i+3)$, $d_{\alpha N}(i,i+4)$ and $d_{\alpha\beta}(i,i+3)$ (see Figure 29 and Section 3.4). In addition, the hydrogen bonds of backbone-amide protons were estimated based on this pattern of ^{1}H signals, which is characteristic for alpha-helical regions. The solved structure includes 81 residues starting with amino-acid residue (Arg) which is the 19th residue in the fusion protein but the first in the native caenopore-5 sequence (see Section 2.1 and Figure 29).

The final 20 conformers which showed the lowest target functions out of 400 calculated structures were superimposed to represent the structure of caenopore-5 (see Figure 30 A). Each structure was calculated independently based on 3044 distance constraints. The statistic data of the structure ensemble are displayed in Table 2. The backbone atoms of the superimposed structures are quite well determined with a root mean square deviation (rmsd) of 0.73 Å for backbone and 0.48 Å for the alpha-helical regions. The 72 % of the Φ and Ψ torsion angles along the backbone are in the most preferred region of the Ramachandran plot while only 1 % are located in disallowed regions. The structure ensemble was used to calculate the average structure of caenopore-5, which is depicted by the ribbon representation (see Figure 30 B).

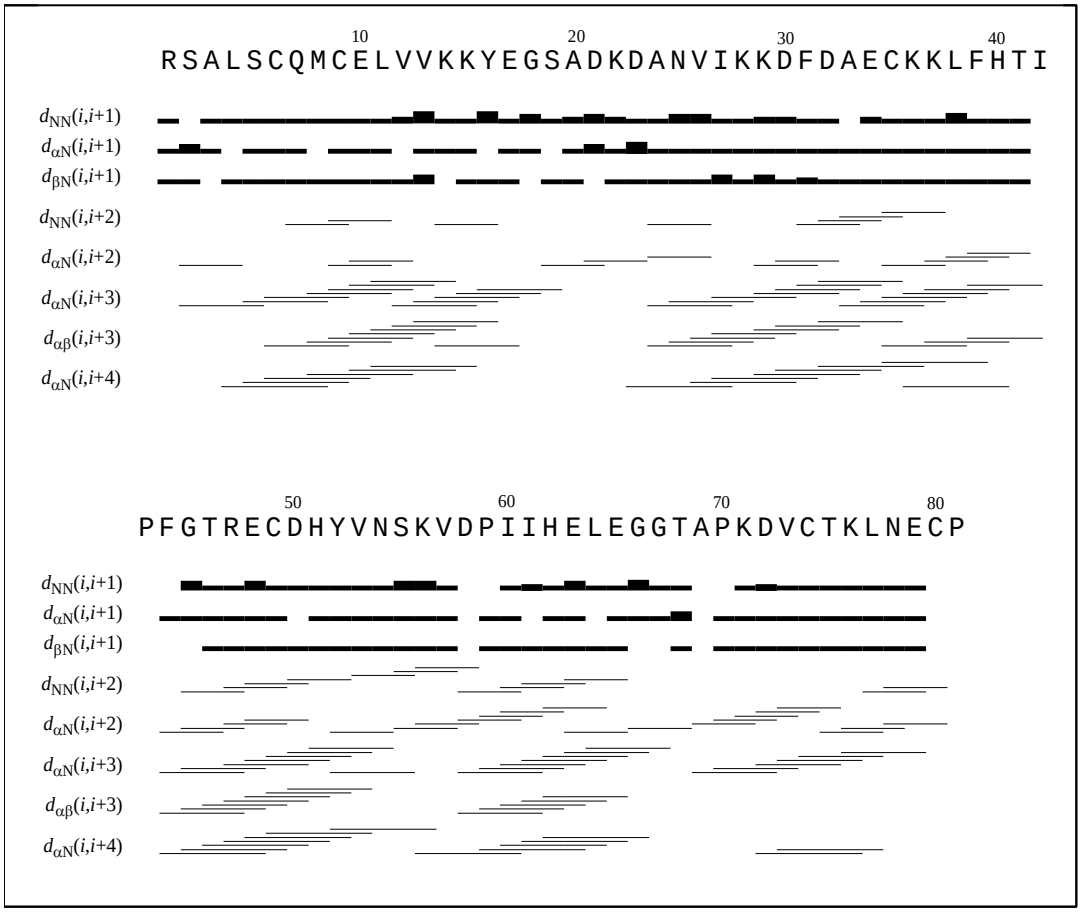


Figure 29: *Sequential and medium-range distances derived from the ^{15}N-edited-NOESY and ^{13}C-edited-NOESY spectra which were used to identify the secondary structural elements. The amino-acid sequence of caenopore-5 is depicted at the top. The bars (lines) connect residues between ^{1}H-^{1}H distance constraints are present. The thickness of the bars in the sequential distances ($d_{NN}(i,i+1)$, $d_{\alpha N}(i,i+1)$, $d_{\beta N}(i,i+1)$) corresponds to the intensities of the ^{1}H resonance signals subdivided into strong, medium and weak. The medium-range distances ($d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+2)$, $d_{\alpha N}(i,i+3)$, $d_{\alpha\beta}(i,i+3)$, $d_{\alpha N}(i,i+4)$) are indicated by thin lines.*

Table 2: Structural statistics for the 20 structures of caenopore-5 with no distance constraint violated by more than 0.5 Å in any structure; rmsd=root mean square deviation.

Distance restraints	
Intraresidual (i – j=0)	1526
Sequential (|i – j|=1)	494
Medium range (2≤|i - j|≤4)	597
Long range (|i - j|≥5)	339
Hydrogen bonds	70
Disulphide bonds	18
All	3044
Pairwise rmsd for residues 2 – 81 in Å	
Mean global backbone rmsd	0.73 ± 0.13
Mean global heavy atom rmsd	1.36 ± 0.14
Pairwise rmsd for secondary structure in Å	
Mean global backbone rmsd	0.48 ± 0.09
Mean global heavy atom rmsd	1.09 ± 0.11
Ramachandran plot (mean values for all 20 conformers)	
Most favored regions (%)	72
Additional allowed regions (%)	23
Generously allowed regions (%)	4
Disallowed regions (%)	1

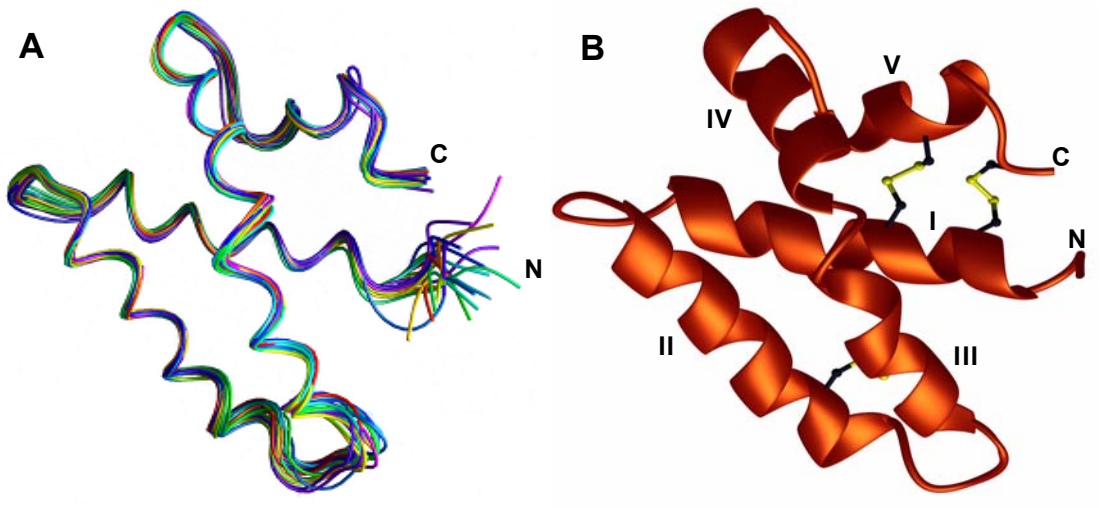


Figure 30: *The structure of caenopore-5. A superposition of backbone atoms in 20 conformers of caenopore-5* ***(A)****. A ribbon representation of the average structure with disulphide bridges shown in yellow* ***(B)****. N and C denote N and C-termini, respectively, numbers from I-V label the helices.*

Caenopore-5 consists of five amphipatic alpha helices including the following amino-acid residues: 5 - 18 (helix-I), 24 - 37 (helix-II), 45 - 53 (helix-III), 56 - 62 (helix-IV) and 72 - 77 (helix-V) (see Figure 30 B and 32 C). The alpha-helical regions are mostly interrupted by short loops or a kink located between helix III and IV. They are all arranged into the tertiary structure which resembles a folded leaf with two halves; the first half is composed of helix I and the shortest helix V and the other half contains helix II, III and IV. This fold is stabilised by three disulphide bridges made of six conserved cysteines: Cys_6-Cys_{80}, Cys_9-Cys_{74} and Cys_{35}-Cys_{49}. The first two disulphide bonds are within the first half, where the last one connects helix II and III within the second half of the leaf. The two halves overlap only partially at the bottom of the bend of the folded leaf as they are mostly pointing into opposite directions relative to each other. The second half is arched due to the kink at Asn 54 (Φ=-137.5°, Ψ=+69.4°) and Ser 55 (Φ=+165.9°, Ψ=-29.5°). In this arrangement, helices II, III and IV of the second half are twisted about 45° compared to helix I and the conformation is stabilised by hydrophobic interactions between amino-acid side chains in helix I with side chains of amino acids from the second half. Helix V of the first half is almost perpendicular to helix IV, which is mainly provoked by contacts between the side chain of Leu 77 and the side chains of Pro 59, His 62 and Glu 63.

3.6.1 Conformers of caenopore-5

The investigation of the spectra during the sequential assignment procedure revealed the presence of residues with double backbone-amide chemical shifts, indicating the existence of two conformers of caenopore-5. To examine the differences between these two forms, the structure of the second conformer was calculated on the basis of ^1H-^1H distance constraints derived from the second set of heterogeneous residues. The structure calculation with of the second set of constraints (conformer II) was performed as described above for the first set of constraints (conformer I). The superposition of both averaged structures (conformer I and II) showed that their backbones are congruent with

very low rmsd over CA equal to 0.3 Å (see Figure 31 A). This indicates that both structures represent the global fold of caenopore-5. These two conformers can be only discriminated on basis of isomerisation of Pro 81 residue (99 residue in fusion protein) (see Figure 31 B, C). The Pro 81 adopts either cis or trans conformation in conformer I or II, respectively.

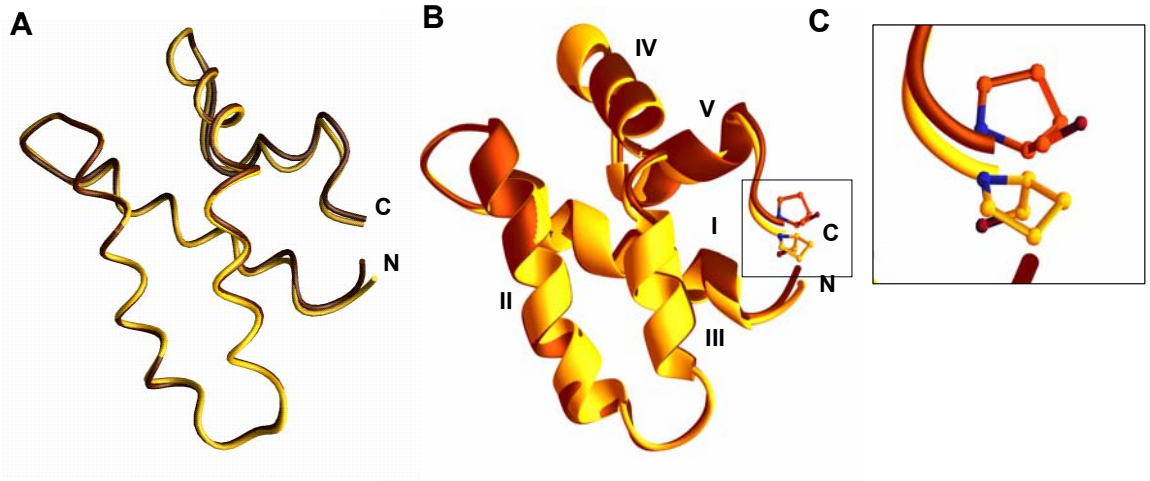


Figure 31: *Superposition of the two averaged structures of caenopore-5 corresponding to conformer **I** (brown) and conformer **II** (yellow) **(A)**. Superposition of the two averaged structures of conformer **I** (brown) and **II** (yellow) with the Pro 81 residue shown in the box **(B).** The close-up of Pro 81 residue present in cis or trans conformation in conformer **I** or **II**, respectively **(C)**. N and C denote N and C-termini, respectively. The helices are numbered from I-V.*

3.6.2 Characterisation of the surface charge distribution

Caenopore-5 has slightly acidic properties with an isoelectric point calculated to be 6.3 and with all charged residues exposed to the solvent. The analysis of the surface electrostatic potential indicates an even charge distribution, except for the uncharged patch prevalent in N and C-terminal regions and the region within the second half of the leaf, which is partially unmasked (see Figure 32 A). The uncharged area is mostly covered with hydrophobic amino acids (see Figure 32 B).

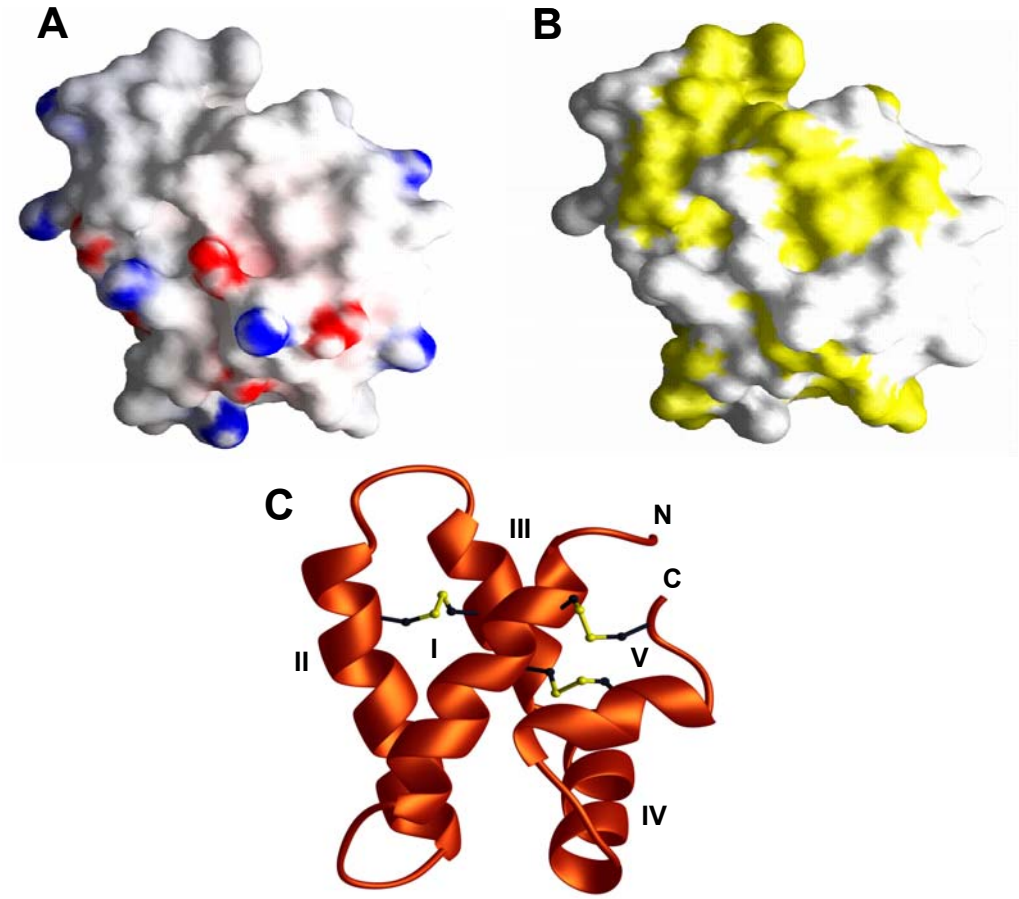


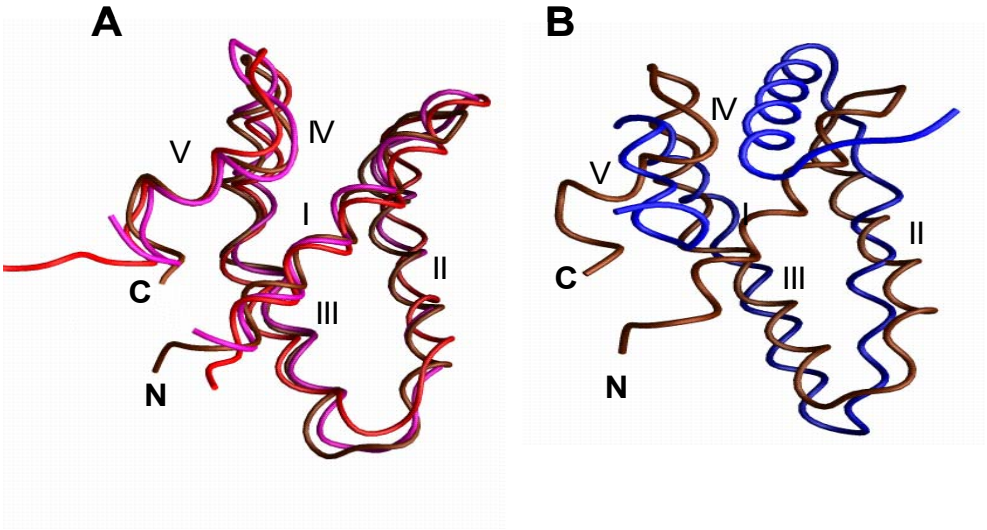
Figure 32: *The structure of caenopore-5. The electrostatic potential of the molecule surface with positively charged regions in blue and negatively charged regions in red* ***(A)****. A distribution of hydrophobic amino-acid residues at the molecule's surface is represented in yellow* ***(B)****. A ribbon representation of the average structure with disulphide bridges shown in yellow in the same orientation as Figure A and B* ***(C)****. N and C denote N and C-termini. The helices are labelled from I-V.*

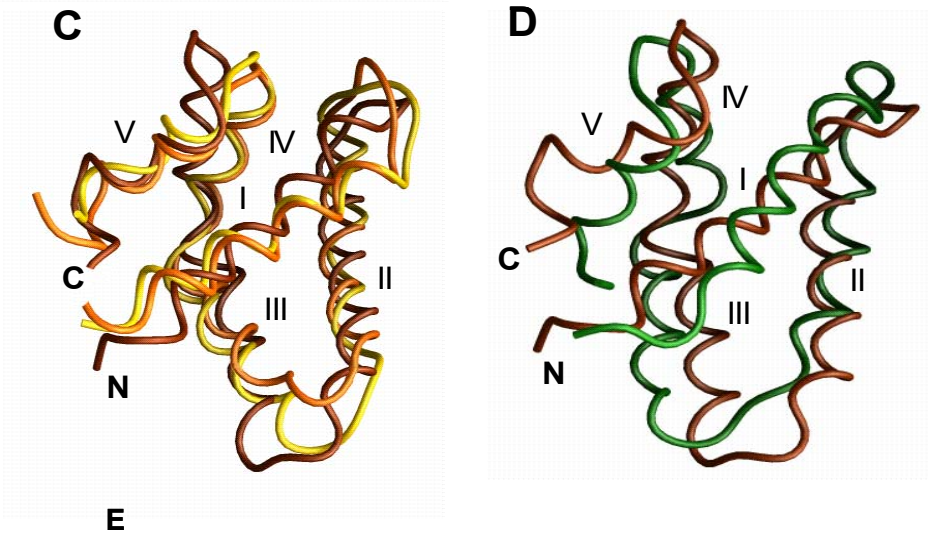
3.7 Structural comparison of caenopore-5 with other members of the SAPLIP family

The three-dimensional structure of caenopore-5 was compared with the structures of NK-lysin, granulysin, amoebapore A, saposin A, saposin B and saposin C. All of those share the characteristic fold of a bundle of amphipathic-alpha helices. Usually five alpha-helical regions (I to V) are distinguished. However, in saposin A, B and C, the single helices III and IV are considered as one helical stretch (helix III) interrupted by a kink [103], [104]. For the purpose of a detailed structural analysis, a superposition of caenopore-5 with all above-mentioned members of the SAPLIP family was performed. All backbone CA atoms of the regular secondary structural elements (helices) were used to calculate the root mean square deviation (see Figure 33 E).

The superposition of both saposin A and C on caenopore-5 revealed the lowest rmsd value of 1.57 Å and 2.5 Å, respectively (see Figure 33 E). The molecules showed a mostly congruent arrangement of the secondary structural elements (see Figure 33 A). Only a small deviation of helix II of saposin C could be observed in comparison to the same region of caenopore-5. Other marginal differences were observed in a distinct reorganisation of the loop regions, especially between helix IV and V in caenopore-5 or between the corresponding helices III and IV in saposin A and C. Additionally, the loop between helix II and III in saposin C slightly pointed towards helix I in comparison to caenopore-5 and saposin A.

In contrast to saposin A, C and caenopore-5, saposin B exists as a homodimer [104]. The fold of the saposin B monomer represents a completely different arrangement of alpha helices, which seems to resemble the structure of saposin C when solved in presence of sodium dodecyl sulphate (SDS) [105]. The comparison of the saposin B monomer with caenopore-5 highlighted these dissimilarities resulting in a high rmsd of 7.1 Å over backbone atoms for alpha-helical regions (see Figure 33 B, E). In the saposin B monomer, the two halves of the leaf are opened in contrast to the compact conformation of caenopore-5.





E

Peptide	PDB ID	rmsd
Saposin A	2DOB.pdb [103]	1.57 Å
Saposin B	1N69.pdb [104]	7.10 Å
Saposin C	1M12.pdb [106]	2.49 Å
NK-lysin	1NKL.pdb [29]	3.19 Å
Granulysin	1L9L.pdb [107]	3.21 Å
Amoebapore A	1OF9.pdb [60]	3.30 Å

Figure 33: *Superposition of caenopore-5 (brown) and other members of the SAPLIP family: saposin A (pink) and saposin C (red)* ***(A)****; saposin B (blue)* ***(B)****; NK-lysin (orange) and granulysin (yellow)* ***(C)****; amoebapore A (green)* ***(D)****.* N and C denote N and C-termini, respectively; numbers from I-V correspond to the order of the alpha helices in caenopore-5. *The root mean square deviations (rmsd) of superimposed structures on caenopore-5 fold are listed* ***(E)****.*

The comparison of the structures of caenopore-5 with porcine NK-lysin and its human counterpart granulysin showed many differences in the arrangement of the alpha-helical regions (see Figure 33 C). Helices III and IV of caenopore-5 were oriented in a different manner in comparison to the corresponding regions in NK-lysin and granulysin. Additionally, helices I and II of NK-lysin and granulysin deviated from corresponding helical regions of caenopore-5. Other differences were observed in a distinct reorganisation of the loop regions; especially the tip of the second half of the leaf of NK-lysin was much more bended toward helix I in comparison to caenopore-5.

The most striking feature of the tertiary fold of amoebapore A, as compared to caenopore-5 was the different orientation of the two halves of the leaf relative to each other (see Figure 33 D). The two halves of the leaf in amoebapore A were almost aligned parallel to each other, whereas they were rotated by approximately 45 ° in the caenopore-5 structure. As a result, not all alpha-helical regions were exactly congruent, which resulted in an rmsd of 3.3 Å for these two proteins (see Figure 33 E).

The structural comparison of caenopore-5 with the structures of other members of the SAPLIP family was additionally supported by the analysis of the electrostatic properties of the solvent-accessible surfaces of these molecules. As described above (see Section 3.6.2) caenopore-5 revealed an even charge distribution, in contrast to the negatively charged surfaces of saposin A (see Figure 34 A, B) and saposin C (data not shown). The isoelectric points of saposin A and C are 4.3 and 4.4, respectively. In contrast to saposin A and C the molecular surfaces of NK-lysin and granulysin are dominated by positively charged residues, which correspond to their high isoelectric point of 8.9 and 10.8, respectively (see Figure 34 C, data only shown for NK-lysin). The molecular surface of amoebapore A is covered with evenly distributed charged residues with only one hydrophobic region, spanning helices I and II (see Figure 34 D).

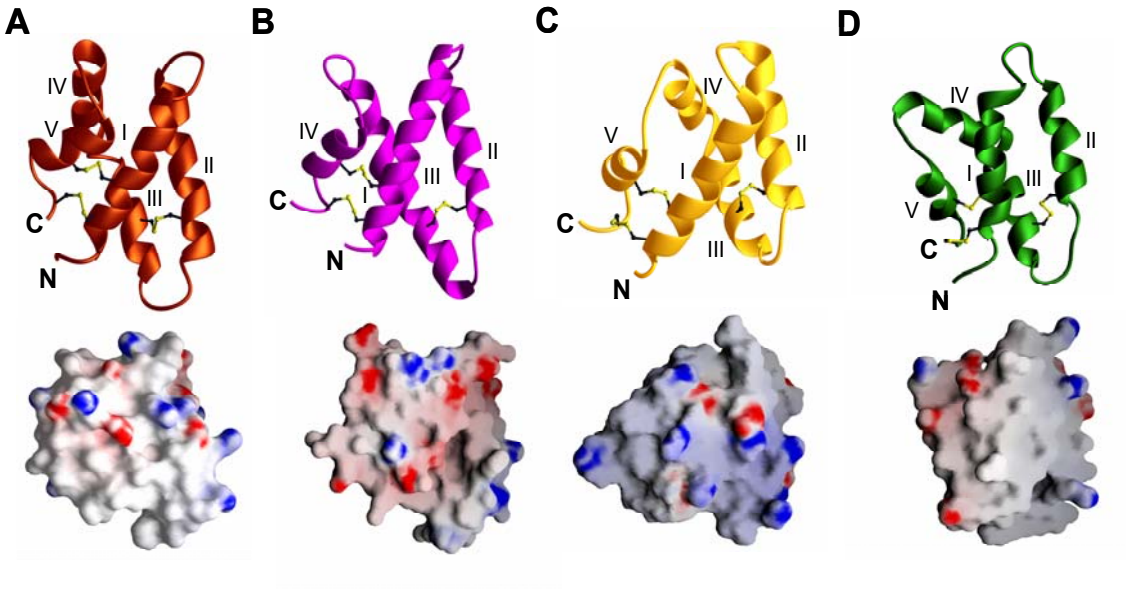


Figure 34: *The ribbon representations of the structures of caenopore-5* ***(A)****, saposin A* ***(B)****, NK-lysin* ***(C)*** *and amoebapore A* ***(D)*** *(upper section) with the corresponding electrostatic potential of their molecular surfaces (lower section), depicted in the same orientation. Disulphide bridges are shown in yellow; N and C denote N and C-termini, respectively; numbers from I-V label the helices; positively and negatively charged regions are represented in blue and red, respectively.*

3.8 Effect of different phospholipids on caenopore-5 conformation

The interaction of caenopore-5 with liposomes composed of various phospholipids was examined by applying far-UV CD-spectroscopy. The N-terminal $(His)_6$ tag in caenopore-5 fusion protein was removed by the factor Xa cleavage to exclude its effect on liposome binding. The correct molecular mass of the cleavage product (9110 Da) was confirmed by MALDI-TOF spectrometry. Conformational changes of caenopore-5 were induced only upon binding to phosphatidylglycerol liposomes (see Figure 35). This interaction could only be observed at a pH below 6. The perturbation of the alpha-helical spectrum of caenopore-5 resulted in a significant decrease in signal intensity with a slight

shift of two minima in the direction of higher wavelengths (208→211 nm, 222→225 nm). Other phospholipids did not induce any conformational changes in caenopore-5 (see Figure 35).

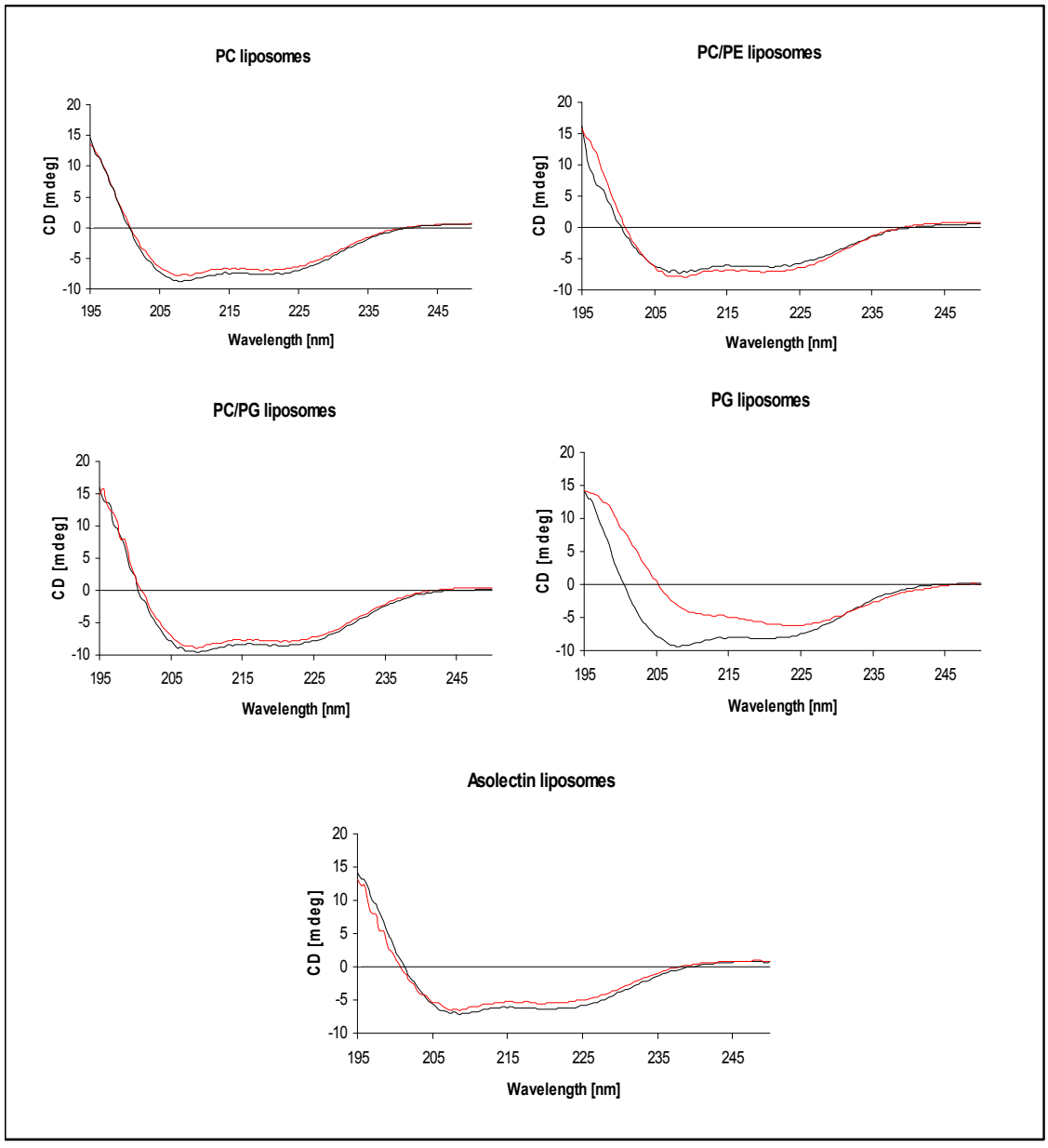


Figure 35: *Interaction of caenopore-5 with various liposomes; PC-phosphatidylcholine; PG-phosphatidylglycerol; PE-phosphatidylethanolamine. Far-UV CD-spectra of caenopore-5 were recorded in 50 mM sodium phosphate at pH=5.2, at 24 °C. Black line: CD-spectrum before mixing with liposomes; red line: CD-spectrum after mixing with liposomes.*

4 Discussion

The soil-living invertebrate *C.elegans* encounters a variety of microorganisms in its natural habitat, which is a great challenge for this simple organism as many of those are not only the essential source of nutrition but also represent potential pathogens. The ability of the worm to survive in the environment with the ubiquitously occurring microbes is inseparably associated with the elaboration of many adaptations. Although various studies were performed on the worm's immunity, many interactions at molecular level still require further research. Caenopore-5 encoded by spp-5 was originally identified as a representative of the SAPLIP family by analysis of the *C.elegans* genome using the consensus sequence method [8]. It is one of many caenopores present in *C.elegans.* It is assumed that they are involved in worm's innate immunity on basis of functional data and sequence homology to peptides of the SAPLIP family, which are known to have antimicrobial activities [9], [14]. The knowledge about the biological role of caenopores is still not complete, but interestingly the inhibition of spp genes activity showed the indispensable role of caenopore-5 in the elimination of bacteria ingested by the worm [9]. Therefore, the structural characterisation of caenopore-5 and comparison to other known folds of members of the SAPLIP family can provide insights in the mode of its interaction with microorganisms.

The determination of the three dimensional structure of caenopore-5 was achieved by means of NMR spectroscopy. Unfortunately, this method does not provide direct structural information in spite of the fact that signals derived from atoms are separated dependent on the diversity of their chemical shifts primarily resulting from conformational impact. Therefore, a prerequisite in the structure determination by NMR spectroscopy is the assignment of the signals to the corresponding atoms in the protein sequence, whose spatial contact examination will define the protein structure. The sequential assignment of caenopore-5 unexpectedly showed that about 20 residues are represented by two different chemical shifts of their backbone amide protons, which suggest the presence of two distinct conformations. The examination of the structures

corresponding to these two conformations of caenopore-5 confirmed the same global fold with no perturbations in the alpha-helical regions or their arrangement, particularly in the affected regions. Since the chemical-shift variations of backbone amide protons are mainly associated with residues in the N and C-terminal part of caenopore-5 and in the fragment spanning the residue range from 60 to 66, subtle local conformational changes can be responsible for this heterogeneity. The analysis of these duplicated residues in the ^{15}N-edited NOESY spectrum of caenopore-5 revealed similar ^{1}H resonances for both conformers, with only slight differences in their intensities, which indicate local perturbations in the structure. Furthermore, amide protons are considered to be very sensitive to even small changes in their chemical environment [108]. Similar variations of the structure were reported for many proteins which often result from disulphide bridge or proline isomerisation [109], [100], [102]. Interestingly, the experimental data demonstrated that the amid-bond between Cys 98 and Pro 99 of caenopore-5 adopts both cis and trans conformation depending on the analysed conformer which explains the presence of affected residues in the C-terminal region (90 - 98). One consequence of the isomerisation of Pro 99 can be the large chemical shift diversity observed in the N-terminal region (19 - 27), as these two regions are in the close proximity resulting from the two disulphide bridges between Cys 98 and Cys 24 and Cys 27 and Cys 92 (see Figure 36). The possibility of the changes in the chirality of the disulphide bridges cannot be excluded as an additional factor contributing to the occurrence of the two forms of caenopore-5. However, the evidence of the disulphide bridge isomerisation resulting from the change of the χ_1 dihedral angle of the cysteins could not be demonstrated due to the lack of direct spectral information in heteronuclear spectra of caenopore-5. In the example of the disulphide bridge isomerisation shown in Figure 37, the conformational changes affect the HB coordinates of one of Cys residue (Cys 38 in BPTI protein) which results only in slight alterations in their distances in relation to surrounding atoms.

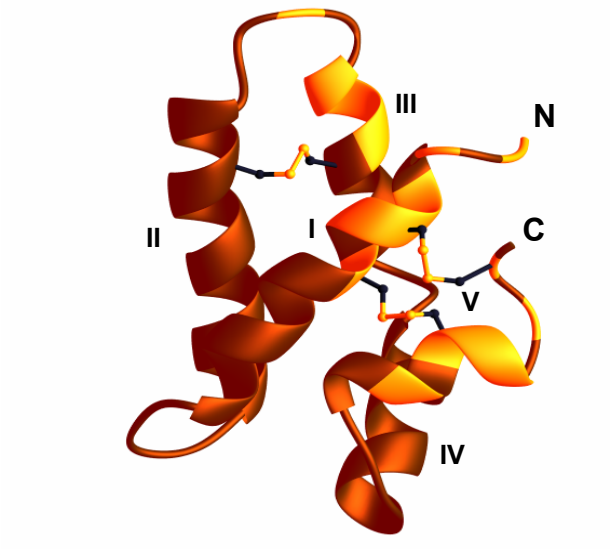


Figure 36: *The structure of caenopore-5. A ribbon representation of the average structure with regions containing duplicated resonances shown in yellow. N and C denote N and C-termini, respectively, numbers from I-V label the helices.*

Interestingly, similar conformational heterogeneity has also been observed in saposin A and even much more pronounced in saposin D (see Table 3) [110]. Noticeable is the fact that the same fragment, residue range from 63 to 66 in caenopore-5, was also affected both in saposin A and saposin D (see Table 3). Although the isomerisation of two conserved proline residues present in the proximity of this amino-acid stretch in saposin A and D was proposed, no experimental data were presented to support this conclusion [110]. Isomerisation of the Pro 61 in caenopore-5 can be excluded as the trans conformation was found in both forms. Alternatively, the disulphide bond chirality should be also taken into consideration due to the close proximity of the conservative Cys residue in caenopore-5 and both in saposin A and D (see Table 3). Another interesting explanation arises from the fact that the N-terminal fragment of caenopore-5 (19 - 27) may affect the region (60 - 66) as they are in close proximity in the tertiary structure (see Figure 36).

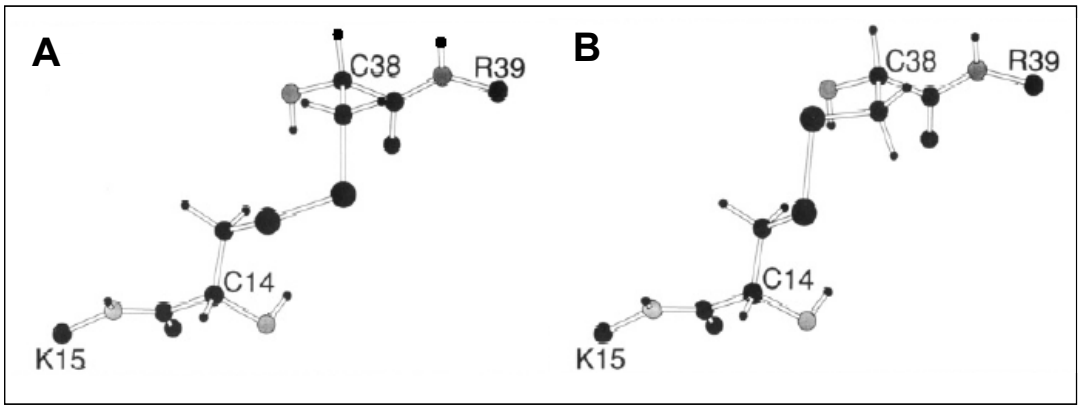


Figure 37: *Model representation of two conformations of the disulphide bridge between C 14 and C 38 in BPTI published by Otting [109] (used with permission). Right-handed* ***(A)*** *and left-handed chirality of disulphide bridge* ***(B)*** *generated by a rotation of –120° around the CA and CB bond (χ_1) of C 38. Carbon atoms are indicated by black balls; nitrogen atoms are indicated by light-shaded balls; sulphur atoms are indicated by big and black balls; hydrogen atoms are indicated by small black balls.*

```
          20                  30                  40                  50
Ca-5    RSALSCQMCELVVKKYEGSADKDA...NVIKKDFDAECKKLFHT
SaA      GSLPCDICKDVVTAAGDMLKDNAT.EEEILVYLEKTCDWLPKP
SaD        GFCEVCKKLVGYLDRNLEKNST.KQEILAALEKGCSFLPDP

          60                  70                  80                  90
Ca-5    IPFGTRECDHYVNSKVDPIIHELEGGTA.PKDVCTKLNECP
SaA     ..NMSASCKEIVDSYLPVILDIIKGEMSRPGEVCSALNLCES
SaD     ...YQKQCDQFVAEYEPVLIEILVEVMD.PSFVCLKIGACPS
```

Table 3**:** *Alignment of amino-acid sequences of caenopore-5, saposin A and D. The residue numbering corresponds to caenopore-5 fused with the N-terminal tag. The conserved cysteines are shown in bold. Proline residues discussed in the text are shown in blue. All duplicated residues are labelled in red. Chemical-shift variations of saposin A and D were described by John [110].*

The solved structure of caenopore-5 confirmed the characteristic saposin-like fold with three disulphide bridges maintaining their structural integrity. In spite of the close overall structural similarity of peptides of the SAPLIP family, a detailed comparison revealed many local differences in the arrangement of the helices and the connecting loop regions between caenopore-5 and these peptides. Furthermore, the inspection of the surface-exposed residues revealed diversity in the charge distribution among the known structures of the members of the SAPLP family. These dissimilarities in the tertiary fold and particularly the variety of the electrostatic surfaces are considered to correlate with their different lipid-associated activities [111], [107], [104], [60]. Interestingly, caenopore-5 mainly resembles the structures of human saposin A and C. Despite the structural similarity, the surface charge pattern of caenopore-5 and saposin A and C varies. Whereas caenopore-5 shows a slightly acid isoelectric point (pI=6.8) with a uniform distribution of positive and negative residues on the surface of the molecule, saposin A and C have a predominantly negatively charged surface. In contrast to these two saposins, saposin B shows the highest deviation from the caenopore-5 fold among all compared structures. However, the open conformation of saposin B is unique at tertiary level among all known structures of the SAPLIP family. In addition, saposins and caenopore-5 diverge in their biological functions. Whereas the essential function of saposins is the stimulation of phospholipids degradation in the lysosomal compartment, caenopore-5 possesses antimicrobial activity and is able to perforate the bacteria membranes [32], [112], [113], [19]. Furthermore, caenopore-5 is only expressed in epithelial cells of the intestine tract and the presence of a signal peptide probably leads to its secretion into the gut lumen, where it can directly act on bacteria. Nevertheless, a comparison of the caenopore-5 structure with the structures of amoebapore A, NK-lysin and granulysin, which also exert antimicrobial activity, showed many deviations in the arrangement of the alpha-helical stretches and loop regions. Despite the conformational differences, amoebapore A with its uniform surface-charge distribution resembles caenopore-5, but both molecules seem to have different mechanism of action.

The antimicrobial activity of amoebapore A is pH-dependent, as the peptide at low pH forms stable dimers, which is the prerequisite for pore formation [60]. Caenopore-5 also displays antimicrobial activity at a pH below 6, but does not generate stable pores in the membrane (personal communication Prof. Dr. M. Leippe). Moreover, a monomeric state of caenopore-5 was revealed at acid pH by means of size-exclusion chromatography. Both NK-lysin and granulysin permeabilise membranes as monomers. The most striking discrepancy between them and caenopore-5, however, is their very high positively charged surface. In general, the positive charges are undoubtedly important for the initial interaction of many antimicrobial peptides with negatively charged membranes of bacteria [114], [115]. However these strong cationic peptides were thought to act on membranes in a carpet-like mechanism with emphasis on the remarkable role of their positive charges in membrane destabilisation [116].

The investigation of the interaction of caenopore-5 with liposomes composed of different lipids showed that the negatively-charged vesicles, comprising only phosphatidylglycerol, induce the conformational change in caenopore-5. This interaction was observed at a pH below 6, which correlates with its antibacterial activity. Since caenopore-5 does not display any significant charged epitopes at its surface, the dependence of these activities on low pH is most likely associated with the partial protonation of negatively charged residues, which leads to an increase of the overall positive charge. The titration studies of Glu residues at the surface of saposin C reported elevated pK values about 5.5 for these amino acids [106]. The protonation of a few His residues present in caenopore-5 (pK is about 6) might also contribute to this effect.

In contrast to negatively charged vesicles, the mixed liposomes composed of zwitterionic lipids, phosphatidylcholine and phosphatidylethanolamine, did not affect the caenopore-5 conformation. Although the binding of caenopore-5 to zwitterionic lipids can not be excluded, the interaction with highly negatively charged phosphatidylglycerol vesicles, imitating bacterial membrane, indicates the particular significance of the electrostatic effect, which triggers the conformational change.

It is also noteworthy that eventual changes of phosphatidylglycerol head groups charge can be neglected as under experimental conditions used (pH=5.2), these are fully deprotonated [117].

The observed conformational alteration of caenopore-5 upon interaction with phosphatidylglycerol liposomes is associated with the remarkable decrease in CD-signal intensity along with the shift of the two minima to longer wavelengths which suggests the integration of the peptide into the hydrophobic environment of the membrane bilayer [118], [119]. Additionally, a simultaneous perturbation of the alpha-helical conformation of caenopore-5 cannot be ruled out. Interestingly, saposin C is thought to reorient its functional helical domains upon interaction with phosphatidylserine vesicles [111]. In contrast to saposin C, NK-lysin seems to preserve its fold upon binding with membrane in spite of the fact that it embeds in the lipid bilayer [86], [120].

The lack of conformational changes of caenopore-5 when it interacts with asolection (mixture of different phospholipids obtained from soybean) and mixed liposomes consisting of phosphatidylcholine and phosphatidylglycerol (3:1) in comparison to phosphatidylglycerol vesicles can be explained by the presence of rare and scattered positively charged residues at its surface, which need a highly dense negative area to trigger the initial interaction with the membrane. Notably, NK-lysin also showed strong preference for negatively charged phospholipids as the addition of phosphatidylglycerol to phosphatidylcholine liposomes (1:3) remarkably increased its pore-forming potential [53]. However, the binding of NK-lysin to the mixed vesicles in comparison to caenopore-5 can be elucidated by the existence of highly positive areas, which facilitate this interaction with the negative charges present even at low density.

The three-dimensional structure of caenopore-5 and the characteristic of its charge statues revealed many interesting features of this peptide which can provide a more complete understanding of biochemical and structural aspects of members of the SAPLIP family, including their specificity against various lipids and mechanism of interaction with these.

5 References

1. Riddle, D. L., Blumenthal, Z., Meyer, B. J., Priess, J. R.: Introduction to *C.elegans*. Plainview: Cold Spring Harbor Laboratory Press,1997: p. 1-22.

2. Hodgkin, J., Kuwabara P. E., Corneliussen, B.: A novel bacterial pathogen, Microbacterium nematophilum, induces morphological change in the nematode *C. elegans*. Curr Biol, 2000, 10 (24): p. 1615-8.

3. Borgonie, G., Claeys, M., Vanfletteren, J., De Waele, D., Coomans, A.: Presence of peritrophic-like membranes in the intestine of three bacteriophagous nematodes (Nematoda: Rhabditida). Fund. Appl. Nematol., 1995, 18: p. 227-233.

4. Pujol, N., Link, E. M., Liu, L. X., Kurz, C. L., Alloing, G., Tan, M. W., Ray, K. P., Solari, R., Johnson, C. D., Ewbank, J. J.: A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. Curr Biol, 2001, 11 (11): p. 809-21.

5. Mallo, G. V., Kurz, C. L., Pujol, N., Granjeaud, S., Kohara, Y., Ewbank, J. J.: Inducible antibacterial defense system in *C. elegans*. Curr Biol, 2002, 12 (14): p. 1209-14.

6. Kato, Y., Aizawa, T., Hoshino, H., Kawano, K., Nitta, K., Zhang, H.:

   abf-1 and abf-2, ASABF-type antimicrobial peptide genes in *Caenorhabditis elegans*. Biochem J, 2002, 361 (Pt 2): p. 221-30.

7. Couillault, C., Pujol, N., Reboul, J., Sabatier, L., Guichou, J. F., Kohara, Y., Ewbank, J. J.: TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. Nat Immunol, 2004, 5 (5): p. 488-94.

8. Banyai, L., Patthy L.: Amoebapore homologs of *Caenorhabditis elegans*. Biochim Biophys Acta, 1998, 1429 (1): p. 259-64.

9. Roeder, T., Leippe, M.: Abstracts of the 21st Annual Meeting of the German Parasitology Association. 17-20 March 2004, Wurzburg, Germany. Int J Med Microbiol, 2004, 293 Suppl 38: p. 44-117.

10. Mochii, M., Yoshida, S., Morita, K., Kohara, Y., Ueno, N.: Identification of transforming growth factor-beta- regulated genes in *caenorhabditis elegans* by differential hybridization of arrayed cDNAs. Proc Natl Acad Sci U S A, 1999, 96 (26): p. 15020-5.

11. Murphy, C. T., McCarroll, S. A, Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H., Kenyon, C.: Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. Nature, 2003, 424 (6946): p. 277-83.

12. Aballay, A., Drenkard, E., Hilbun, L. R., Ausubel, F. M.: *Caenorhabditis elegans* innate immune response triggered by *Salmonella enterica* requires intact LPS and is mediated by a MAPK signaling pathway. Curr Biol, 2003, 13 (1): p. 47-52.

13. Levashina, E. A., Moita, L. F., Blandin, S., Vriend, G., Lagueux, M., Kafatos, F. C.: Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. Cell, 2001, 104 (5): p. 709-18.

14. Schulenburg, H., Kurz, C. L., Ewbank, J. J.: Evolution of the innate immune system: the worm perspective. Immunol Rev, 2004, 198: p. 36-58.

15. Couillault, C., Ewbank,J. J.: Diverse bacteria are pathogens of *Caenorhabditis elegans*. Infect Immun, 2002, 70 (8): p. 4705-7.

16. Mylonakis, E., Ausubel, F. M., Perfect, J. R., Heitman, J., Calderwood, S. B.: Killing of *Caenorhabditis elegans* by Cryptococcus neoformans as a model of yeast pathogenesis. Proc Natl Acad Sci U S A, 2002, 99 (24): p. 15675-80.

17. Garsin, D. A., Sifri, C. D., Mylonakis, E., Qin, X., Singh, K. V., Murray, B. E., Calderwood, S. B., Ausubel, F. M.: A simple model host for identifying Gram-positive virulence factors. Proc Natl Acad Sci U S A, 2001, 98 (19): p. 10892-7.

18. Kurz, C. L., Tan, M. W.: Regulation of aging and innate immunity in *C. elegans*. Aging Cell, 2004, 3 (4): p. 185-93.

19. Stanisak, M., Roeder T., Leippe M.: Abstract of 10th International Congress; International Society of Developmental and Comparative Immunology, 2006.

20. Tamaru, T., Fujibayashi, S., Brown, W. R., Wenger, D. A.: Immunocytochemical localization of sphingolipid activator protein-1, the sulfatide/GM1 ganglioside activator, to lysosomes in human liver and colon. Histochemistry, 1986, 86 (2): p. 195-200.

21. Paton, B. C., Hughes, J. L., Harzer, K., Poulos, A.: Immunocytochemical localization of sphingolipid activator protein 2 (SAP-2) in normal and SAP-deficient fibroblasts. Eur J Cell Biol, 1990, 51 (1): p. 157-64.

22. O'Brien, J. S., Kishimoto, Y.: Saposin proteins: structure, function, and role in human lysosomal storage disorders. Faseb J, 1991, 5 (3): p. 301-8.

23. Bruhn, H.: A short guided tour through functional and structural features of saposin-like proteins. Biochem J, 2005, 389 (Pt 2): p. 249-57.

24. Leippe, M., Ebel, S., Schoenberger, O. L., Horstmann, R. D., Muller-Eberhard, H. J.: Pore-forming peptide of pathogenic *Entamoeba histolytica*. Proc Natl Acad Sci U S A, 1991, 88 (17): p. 7659-63.

25. Leippe, M., Tannich, E., Nickel, R., van der Goot, G., Pattus, F., Horstmann, R. D., Muller-Eberhard, H. J.: Primary and secondary structure of the pore-forming peptide of pathogenic *Entamoeba histolytica*. Embo J, 1992, 11 (10): p. 3501-6.

26. Guruprasad, K., Tormakangas, K., Kervinen, J., Blundell, T. L.: Comparative modelling of barley-grain aspartic proteinase: a structural rationale for observed hydrolytic specificity. FEBS Lett, 1994, 352 (2): p. 131-6.

27. Andersson, M., Gunne, H., Agerberth, B., Boman, A., Bergman, T., Sillard, R., Jornvall, H., Mutt, V., Olsson, B., Wigzell, H.: NK-lysin, a novel effector peptide of cytotoxic T and NK cells. Structure and cDNA cloning of the porcine form, induction by interleukin 2, antibacterial and antitumour activity. Embo J, 1995, 14 (8): p. 1615-25.

28. Jongstra, J., Schall, T. J., Dyer, B. J., Clayberger, C., Jorgensen, J., Davis, M. M.,Krensky, A. M.: The isolation and sequence of a novel gene from a human functional T cell line. J Exp Med, 1987, 165 (3): p. 601-14.

29. Liepinsh, E., Andersson, M., Ruysschaert, J., M.Otting, G.: Saposin fold revealed by the NMR structure of NK-lysin. Nat Struct Biol, 1997, 4 (10): p. 793-5.

30. Gonzalez, C., Langdon, G. M., Bruix, M., Galvez, A.,Valdivia, E., Maqueda, M., Rico, M.: Bacteriocin AS-48, a microbial cyclic polypeptide structurally and functionally related to mammalian NK-lysin. Proc Natl Acad Sci U S A, 2000, 97 (21): p. 11221-6.

31. Lee, J. H., Yang, S. T., Rho, S. H., Kim, S. Y., Kim, Y. R., Kim, M. K., Kang, G. B., Kim, J. I., Rhee, J. H., Eom, S. H.: Crystal structure and functional studies reveal that PAS factor from *Vibrio vulnificus* is a novel member of the saposin-fold family. J Mol Biol, 2006, 355 (3): p. 491-500.

32. Ho, M. W., O'Brien, J. S.: Gaucher's disease: deficiency of 'acid' -glucosidase and reconstitution of enzyme activity in vitro. Proc Natl Acad Sci U S A, 1971, 68 (11): p. 2810-3.

33. Morimoto, S., Martin, B. M.,Yamamoto, Y., Kretz, K. A., O'Brien, J. S., Kishimoto, Y.: Saposin A: second cerebrosidase activator protein. Proc Natl Acad Sci U S A, 1989, 86 (9): p. 3389-93.

34. Morimoto, S., Morimoto, S., Martin, B. M., Kishimoto, Y., O'Brien, J. S.: Saposin D: a sphingomyelinase activator. Biochem Biophys Res Commun, 1988, 156 (1): p. 403-10.

35. Vaccaro, A. M., Tatti, M., Ciaffoni, F., Salvioli, R., Maras, B., Barca, A.: Function of saposin C in the reconstitution of glucosylceramidase by phosphatidylserine liposomes. FEBS Lett, 1993, 336 (1): p. 159-62.

36. Ciaffoni, F., Tatti, M., Salvioli, R., Vaccaro, A. M.: Interaction of saposin D with membranes: effect of anionic phospholipids and sphingolipids. Biochem J, 2003, 373 (Pt 3): p. 785-92.

37. Vogel, A., Schwarzmann, G., Sandhoff, K.: Glycosphingolipid specificity of the human sulfatide activator protein. Eur J Biochem, 1991, 200 (2): p. 591-7.

38. Locatelli-Hoops, S., Remmel, N., Klingenstein, R., Breiden, B., Rossocha, M., Schoeniger, M., Koenigs, C., Saenger, W., Sandhoff, K.: Saposin A mobilizes lipids from low cholesterol and high bis(monoacylglycerol)phosphate-containing membranes: patient variant Saposin A lacks lipid extraction capacity. J Biol Chem, 2006, 281 (43): p. 32451-60.

39. Hagen, F. S., Grant, F. J., Kuijper, J. L., Slaughter, C. A., Moomaw, C. R., Orth, K., O'Hara, P. J., Munford, R. S.: Expression and characterization of recombinant human acyloxyacyl hydrolase, a leukocyte enzyme that deacylates bacterial lipopolysaccharides. Biochemistry, 1991, 30 (34): p. 8415-23.

40. Ponting, C.P.: Acid sphingomyelinase possesses a domain homologous to its activator proteins: saposins B and D. Protein Sci, 1994, 3 (2): p. 359-61.

41. Kervinen, J., Tobin, G. J., Costa, J., Waugh, D. S., Wlodawer, A., Zdanov, A.:Crystal structure of plant aspartic proteinase prophytepsin: inactivation and vacuolar targeting. Embo J, 1999, 18 (14): p. 3947-55.

42. Verissimo, P., Faro, C., Moir, A. J., Lin, Y., Tang, J., Pires, E.: Purification, characterization and partial amino acid sequencing of two new aspartic proteinases from fresh flowers of *Cynara cardunculus L*. Eur J Biochem, 1996, 235 (3): p. 762-8.

43. Staab, J. F., Ginkel, D. L., Rosenberg, G. B., Munford, R. S.: A saposin-like domain influences the intracellular localization, stability, and catalytic activity of human acyloxyacyl hydrolase. J Biol Chem, 1994, 269 (38): p. 23736-42.

44. Ida, H., Rennert, O. M., Eto, Y., Chan, W. Y.: Cloning of a human acid sphingomyelinase cDNA with a new mutation that renders the enzyme inactive. J Biochem (Tokyo), 1993, 114 (1): p. 15-20.

45. Tormakangas, K., Hadlington, J. L., Pimpl, P., Hillmer, S., Brandizzi, F., Teeri, T. H., Denecke, J.: A vacuolar sorting domain may also influence the way in which proteins leave the endoplasmic reticulum. Plant Cell, 2001, 13 (9): p. 2021-32.

46. Runeberg-Roos, P., Tormakangas, K., Ostman, A.: Primary structure of a barley-grain aspartic proteinase. A plant aspartic proteinase resembling mammalian cathepsin D. Eur J Biochem, 1991, 202 (3): p. 1021-7.

47. Ramalho-Santos, M., Verissimo, P., Cortes, L., Samyn, B., Van Beeumen, J., Pires, E., Faro, C.: Identification and proteolytic processing of procardosin A. Eur J Biochem, 1998, 255 (1): p. 133-8.

48. Ponting, C. P., Russell, R. B.: Swaposins: circular permutations within genes encoding saposin homologues. Trends Biochem Sci, 1995, 20 (5): p. 179-80.

49. Tanaka, Y., Takei, T., Aiba, T., Masuda, K., Kiuchi, A., Fujiwara, T.: Development of synthetic lung surfactants. J Lipid Res, 1986, 27 (5): p. 475-85.

50. Poulain, F. R., Allen, L., Williams, M. C., Hamilton, R. L., Hawgood, S.: Effects of surfactant apolipoproteins on liposome structure: implications for tubular myelin formation. Am J Physiol, 1992, 262 (6 Pt 1): p. L730-9.

51. Hawgood, S., Derrick, M. Poulain, F.: Structure and properties of surfactant protein B. Biochim Biophys Acta, 1998, 1408 (2-3): p. 150-60.

52. Stenger, S., Hanson, D. A.,Teitelbaum, R., Dewan, P., Niazi, K. R., Froelich, C. J., Ganz, T., Thoma-Uszynski, S., Melian, A., Bogdan, C., Porcelli, S. A., Bloom, B. R., Krensky, A. M., Modlin, R. L.: An antimicrobial activity of cytolytic T cells mediated by granulysin. Science, 1998, 282 (5386): p. 121-5.

53. Bruhn, H., Riekens, B., Berninghausen, O., Leippe, M.: Amoebapores and NK-lysin, members of a class of structurally distinct antimicrobial and cytolytic peptides from protozoa and mammals: a comparative functional analysis. Biochem J, 2003, 375 (Pt 3): p. 737-44.

54. Linde, C. M., Grundstrom, S., Nordling, E., Refai, E., Brennan, P. J., Andersson, M.: Conserved structure and function in the granulysin and NK-lysin peptide family. Infect Immun, 2005, 73 (10): p. 6332-9.

55. Leippe, M., Herbst, R.: Ancient weapons for attack and defense: the pore-forming polypeptides of pathogenic enteric and free-living amoeboid protozoa. J Eukaryot Microbiol, 2004, 51 (5): p. 516-21.

56. Espino, A. M., Hillyer, G. V.: Molecular cloning of a member of the *Fasciola hepatica* saposin-like protein family. J Parasitol, 2003, 89 (3): p. 545-52.

57. Lee, J. Y., Cho, P. Y., Kim, T. Y., Kang, S. Y., Song, K. Y., Hong, S. J.: Hemolytic activity and developmental expression of pore-forming peptide, clonorin. Biochem Biophys Res Commun, 2002, 296 (5): p. 1238-44.

58. Pena, S. V., Hanson, D. A., Carr, B. A., Goralski, T. J., Krensky, A. M.: Processing, subcellular localization, and function of 519 (granulysin), a human late T cell activation molecule with homology to small, lytic, granule proteins. J Immunol, 1997, 158 (6): p. 2680-8.

59. Ma, L. L., Spurrell, J. C., Wang, J. F., Neely, G. G., Epelman, S., Krensky, A. M., Mody, C. H.: CD8 T cell-mediated killing of *Cryptococcus neoformans* requires granulysin and is dependent on CD4 T cells and IL-15. J Immunol, 2002, 169 (10): p. 5787-95.

60. Hecht, O., Van Nuland, N. A., Schleinkofer, K., Dingley, A. J., Bruhn, H., Leippe, M., Grotzinger, J.: Solution structure of the pore-forming protein of *Entamoeba histolytica*. J Biol Chem, 2004, 279 (17): p. 17834-41.

61. Ryan, M. A., Akinbi, H. T., Serrano, A. G., Perez-Gil, J., Wu, H., McCormack, F. X., Weaver, T. E.: Antimicrobial activity of native and synthetic surfactant protein B peptides. J Immunol, 2006, 176 (1): p. 416-25.

62. Egas, C., Lavoura, N., Resende, R., Brito, R. M., Pires, E., de Lima, M. C., Faro, C.: The saposin-like domain of the plant aspartic proteinase precursor is a potent inducer of vesicle leakage. J Biol Chem, 2000, 275 (49): p. 38190-6.

63. Winau, F., Schwierzeck, V., Hurwitz, R., Remmel, N., Sieling, P. A., Modlin, R. L., Porcelli, S. A., Brinkmann, V., Sugita, M., Sandhoff, K., Kaufmann, S. H., Schaible, U. E.: Saposin C is required for lipid presentation by human CD1b. Nat Immunol, 2004, 5 (2): p. 169-74.

64. Zhou, D., Cantu, C., 3rd Sagiv, Y., Schrantz, N., Kulkarni, A. B., Qi, X., Mahuran, D. J., Morales, C. R., Grabowski, G. A., Benlagha, K., Savage, P., Bendelac, A., Teyton, L.: Editing of CD1d-bound lipid antigens by endosomal lipid transfer proteins. Science, 2004, 303 (5657): p. 523-7.

65. Bloch, F., Hansen, W. W., Packard, M.: Nuclear induction. Physical Review, 1946, 69 (3-4): p. 474-485.

66. Purcell, E. M., Torrey, H. C., Pound, R. V.: Resonance Absorption by Nuclear Magnetic Moments in a Solid. Physical Review, 1946, 69 (1-2): p. 37-38.

67. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Skelton N. J.: Protein NMR spectroscopy. Principles and practice. San Diego: Academic Press, 1996.

68. Wüthrich, K.: NMR of Proteins and Nucleic Acids. New York: Wiley, 1986.

69. Ernst, R. R., Anderson, W. A.: Application of Fourier Transform spectroscopy to magnetic resonance. Rev. Sci. Instrum. 1966, 37: p. 93-102.

70. Bodenhausen, G., Freeman, R., Morris, G. A.: A simple pulse sequence for selective excitation in Fourier Transform NMR. J. Magn. Reson., 1976, 23: p. 171 - 175.

71. Cooley, J. W., Tukey, J. W.: An algorithm for the machine calculation of complex Fourier series. Mathematics of Computation, 1965, 19: p. 297-301.

72. Aue, W. P., Bartholdi, E., Ernst R. R.: Two - dimensional spectroscopy, application to nuclear magnetic resonance. J. Chem. Phys., 1976, 64 (5): p. 2229 -2246.

73. Jeener, J., Meier, B. H., Bachmann, P., Ernst, R. R.: Investigation of exchange processes by two-dimensional NMR spectroscopy. J. Chem. Phys., 1979, 71: p. 4546-4553.

74. Ramsey, N. F., Purcell, E. M.: Interactions between Nuclear Spins in Molecules. Phys. Rev., 1952, 85 (1): p. 143–144.

75. Noggle, J. H., Shirmer, R. E.: The Nuclear Overhauser effect:chemical applications. New York: Academic Press, 1971: p. 1- 259.

76. Lottspeich, F., Zorbas, H.: Bioanalytik. Heidelberg-Berlin: Spektrum Akademischer Verlag, 1998.

77. Zuiderweg, E. R., Fesik, S. W.: Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. Biochemistry, 1989, 28 (6): p. 2387-91.

78. Oschkinat, H., Griesinger, C., Kraulis, P. J., Sorensen, O. W., Ernst, R. R., Gronenborn, A. M., Clore, G. M.: Three-dimensional NMR spectroscopy of a protein in solution. Nature, 1988, 332 (6162): p. 374-6.

79. Morris, G. A.: Indirect measurment of proton relaxation rates by 'INEPT' polarization transfer to carbon- 13: proton spin- lattice relaxation in cholesteryl acetate solutions. J. Magn. Reson., 1980, 41: p. 185 -188.

80. Olejniczak, E. T., Xu, R. X., Fesik, S. W.: A 4D HCCH-TOCSY experiment for assigning the side chain 1H and 13C resonances of proteins. J Biomol NMR, 1992, 2 (6): p. 655-9.

81. Shaka, A. J., Keeler, J.: Broadband spin decoupling in isotropic liquids. Progress in NMR Spectroscopy, 1987, 19: p. 47 - 129.

82. Laemmli, U. K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 1970, 227 (5259): p. 680-5.

83. Gill, S. C., von Hippel, P. H.: Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem, 1989, 182 (2): p. 319-26.

84. Kelly, S. M., Jess, T. J., Price, N. C.: How to study proteins by circular dichroism. Biochim Biophys Acta, 2005, 1751 (2): p. 119-39.

85. Chen, G. C., Yang, J. T.: Two-point calibration of circular dichrometer with d-10-camphorsulfonic acid. Anal. Lett., 1977, 10 (14): p. 1195-1207.

86. Ruysschaert, J. M., Goormaghtigh, E., Homble, F., Andersson, M., Liepinsh, E., Otting, G.: Lipid membrane binding of NK-lysin. FEBS Lett, 1998, 425 (2): p. 341-4.

87. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., Bax, A.: NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR., 1995, 6 (3): p. 277-93.

88. Johnson, B. A., Blevins, R. A.: A computer program for the visualisation and analysis of NMR data. J. Biomol NMR., 1994, 4: p. 603-614.

89. Grzesiek, S., Bax, A.: An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. J Magn Reson, 1992, 99: p. 201-209.

90. Montelione, G. T., Lyons, B. A., Emerson, S. D., Tashiro, M.: An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins. Journal of the American Chemical Society, 1992, 114 (27): p. 10974-5.

91. Grzesiek, S., Bax, A.: Correlating Backbone Amide and Side Chain Resonances in Larger Proteins by Multiple Relayed Triple Resonance NMR. J. Am. Chem. Soc., 1992, 114: p. 6291-6293.

92. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., Clore, G. M.: Overcoming the overlap problem in the assignment of 1H NMR spectra of larger proteins by use of three-dimensional heteronuclear 1H-15N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. Biochemistry, 1989, 28 (15): p. 6150-6.

93. Bax, A., Clore, G. M., Gronenborn, A. M.: Proton-proton correlation via isotropic mixing of carbon-13 magnetization, a new three-dimensional approach for assigning proton and carbon-13 spectra of carbon-13-enriched proteins. Journal of Magnetic Resonance, 1990, 88 (2): p. 425-31.

94. Guntert, P., Mumenthaler, C., Wuthrich, K.: Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol, 1997, 273 (1): p. 283-98.

95. Kumar, A., Ernst, R R., Wuthrich, K.: A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. Biochem Biophys Res Commun, 1980, 95 (1): p. 1-6.

96. Kraulis, P. J.: MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. Journal of Applied Crystallography, 1991, 24: p. 946-950.

97. Nicholls, A., Sharp, K. A., Honig, B.: Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins, 1991, 11 (4): p. 281-96.

98. Kraulis, P. J.: Protein three-dimensional structure determination and sequence-specific assignment of 13C and 15N-separated NOE data. A novel real-space ab initio approach. J Mol Biol, 1994, 243 (4): p. 696-718.

99. Wüthrich, K., Billeter, M., Braun, W.: Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. J Mol Biol, 1984, 180 (3): p. 715-40.

100. Chazin, W. J., Kordel, J., Drakenberg, T., Thulin, E., Brodin, P., Grundstrom, T., Forsen, S.: Proline isomerism leads to multiple folded conformations of calbindin D9k: direct evidence from two-dimensional 1H NMR spectroscopy. Proc Natl Acad Sci U S A, 1989, 86 (7): p. 2195-8.

101. Santiveri, C. M., Perez-Canadillas, J. M., Vadivelu, M. K., Allen, M. D., Rutherford, T. J., Watkins, N. A., Bycroft, M.: NMR structure of the alpha-hemoglobin stabilizing protein: insights into conformational heterogeneity and binding. J Biol Chem, 2004, 279 (33): p. 34963-70.

102. Sarkar, P., Reichman, C., Saleh, T., Birge, R. B., Kalodimos, C. G.: Proline cis-trans isomerization controls autoinhibition of a signaling protein. Mol Cell, 2007, 25 (3): p. 413-26.

103. Ahn, V. E., Leyko, P., Alattia, J. R., Chen, L., Prive, G. G.: Crystal structures of saposins A and C. Protein Sci, 2006, 15 (8): p. 1849-57.

104. Ahn, V. E., Faull, K. F., Whitelegge, J. P., Fluharty, A. L., Prive, G. G.: Crystal structure of saposin B reveals a dimeric shell for lipid binding. Proc Natl Acad Sci U S A, 2003, 100 (1): p. 38-43.

105. Hawkins, C. A., de Alba, E., Tjandra, N.: Solution structure of human saposin C in a detergent environment. J Mol Biol, 2005, 346 (5): p. 1381-92.

106. de Alba, E. Weiler, S., Tjandra, N.: Solution structure of human saposin C: pH-dependent interaction with phospholipid vesicles. Biochemistry, 2003, 42 (50): p. 14729-40.

107. Anderson D. H., Sawaya M. R., Cascio, D., Ernst, W., Modlin, R., Krensky, A., Eisenberg, D.: Granulysin crystal structure and a structure-derived lytic mechanism. J Mol Biol, 2003, 325 (2): p. 355-65.

108. Wüthrich, K.: NMR in Biological Research: Peptides and Proteins. Amsterdam: North Holland,1976.

109. Otting, G., Liepinsh, E., Wuthrich, K.: Disulfide bond isomerization in BPTI and BPTI(G36S): an NMR study of correlated mobility in proteins. Biochemistry, 1993, 32 (14): p. 3571-82.

110. John, M., Wendeler, M., Heller, M., Sandhoff, K., Kessler, H.: Characterization of human saposins by NMR spectroscopy. Biochemistry, 2006, 45 (16): p. 5206-16.

111. Qi, X., Grabowski, G. A.: Differential membrane interactions of saposins A and C: implications for the functional specificity. J Biol Chem, 2001, 276 (29): p. 27010-7.

112. Spiegel, R., Bach, G., Sury, V., Mengistu, G., Meidan, B., Shalev, S., Shneor, Y., Mandel, H., Zeigler, M.: A mutation in the saposin A coding region of the prosaposin gene in an infant presenting as Krabbe disease: first report of saposin A deficiency in humans. Mol Genet Metab, 2005, 84 (2): p. 160-6.

113. Kretz, K. A., Carson, G. S., Morimoto, S., Kishimoto, Y., Fluharty, A. L., O'Brien, J. S.: Characterization of a mutation in a family with saposin B deficiency: a glycosylation site defect. Proc Natl Acad Sci U S A, 1990, 87 (7): p. 2541-4.

114. Bessalle, R., Haas, H., Goria, A., Shalit, I., Fridkin, M.: Augmentation of the antibacterial activity of magainin by positive-charge chain extension. Antimicrob Agents Chemother, 1992, 36 (2): p. 313-7.

115. Dathe, M., Wieprecht, T., Nikolenko, H., Handel, L., Maloy, W. L., MacDonald, D. L., Beyermann, M., Bienert, M.: Hydrophobicity, hydrophobic moment and angle subtended by charged residues modulate antibacterial and haemolytic activity of amphipathic helical peptides. FEBS Lett, 1997, 403 (2): p. 208-12.

116. Miteva, M., Andersson, M., Karshikoff, A., Otting, G.: Molecular electroporation: a unifying concept for the description of membrane pore formation by antibacterial peptides, exemplified with NK-lysin. FEBS Lett, 1999, 462 (1-2): p. 155-8.

117. Lakhdar-Ghazal, F., Tichadou, J. L., Tocanne, J. F.: Effect of pH and monovalent cations on the ionization state of phosphatidylglycerol in monolayers. An experimental (surface potential) and theoretical (Gouy-Chapman) approach. Eur J Biochem, 1983, 134 (3): p. 531-7.

118. Ludtke, S. J., He, K., Wu, Y., Huang, H. W.: Cooperative membrane insertion of magainin correlated with its cytolytic activity. Biochim Biophys Acta, 1994, 1190 (1): p. 181-4.

119. Yang, L., Harroun, T. A., Weiss, T. M., Ding, L., Huang, H. W.: Barrel-stave model or toroidal model? A case study on melittin pores. Biophys J, 2001, 81 (3): p. 1475-85.

120. Andersson, M., Andersson, M., Curstedt, T., Jornvall, H., Johansson, J.: An amphipathic helical motif common to tumourolytic polypeptide NK-lysin and pulmonary surfactant polypeptide SP-B. FEBS Lett, 1995, 362 (3): p. 328-32.

Abbreviations

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
A	Alanine
ABF	Antibacterial factor
AI	After induction
Ala	Alanine
AOAH	Acyloxyacyl hydrolase
APS	Ammonium peroxydisulphate
Arg	Arginine
ASM	Acid sphingomyelinase
Asn	Asparagine
BI	Before induction
BPTI	Basic pancreatic trypsin inhibitor
C	Cysteine
C.elegans	*Caenorhabditis elegans*
CA	Carbon alpha
CB	Carbon beta
CD	Circular dichroism
Cys	Cysteine
D	Aspartic acid
DTT	Dithiothreitol
E 1, E 2,…	Elution steps
E	Glutamic acid

E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
F	Phenylalanine
FFT	Fast Fourier Transformation
FID	Free induction decay
FOXO	Forkhead box O
FPLC	Fast protein liquid chromatography
G	Glycine
Gdn · HCl	Guanidine hydrochloride
Gln	Glutamine
Glu	Glutamic acid
GPI	Glycosylphosphatidylinositol
H	Histidine
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid
His	Histidine
HSQC	Heteronuclear single quantum coherence
I	Isoleucine
Ile	Isoleucine
INEPT	Insensitive nuclei enhanced by polarisation transfer
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
K	Lysine
kDa	Kilo Dalton
L	Leucine
LB	Luria Broth
Leu	Leucine
LUV	Large unilamellar vesicles
M	Methionine
MALDI-TOF	Matrix-assisted laser desorption/ionisation-time of flight

MAPK	Mitogen-activated protein kinase
MRW	Mean residue weight
ms	Millisecond
MW	Molecular weight
MWCO	Molecular-weight cutoff
N	Asparagine
Ni^{+2}-NTA	Nickel-nitrilotriacetic acid
NK cells	Natural killer cells
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser and exchange spectroscopy
OD	Optical density
P	Proline
PAMPs	Pathogen-associated molecular patterns
PC	Phosphatidylcholine
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
Phe	Phenylalanine
ppm	Parts per milion
Pro	Proline
PSI	Plant-specific insert
Q	Glutamine
R	Arginine
RF	Radio frequency
rmsd	Root mean square deviation
S	Serine
SAPLIP	Saposin-like protein
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Ser	Serine
SP-B	Surfactant protein B
Spp	Saposin-like protein

T	Threonine
TEMED	N, N, N', N'- tetramethylethylenediamine
TEP	Thioester-containing protein
TGF-β	Transforming growth factor-β
TOCSY	Total correlation spectroscopy
TRIS	Trishydroxymethylaminomethane
UV	Ultraviolet
V	Valine
W	Washing step
Y	Tyrosine

Acknowledgements

I would like to express my sincerest gratitude to my supervisor, Prof. Dr. Joachim Grötzinger for giving me the opportunity to work in his group. I am particularly indebted for his insightful advices, continuous support and encouragement throughout the course of this research.

I would like to thank Prof. Dr. Stefan Rose-John for the warm welcome in the Biochemistry Institute and his advices.

I thank Prof. Dr. Matthias Leippe for introducing me into the caenopore-5 topic, his valuable comments and reviewing my thesis.

I would like to express my gratitude to Prof. Dr Ulrich Zähringer for the possibility to perform NMR experiments at the Reaserch Center in Borstel.

I wish to thank Dr. Andrew Dingley, Dr. Inken Lorenzen and Heiko Käßner for recording the NMR spectra.

I am also indebted to Mareike Stanisak for her help and precious advices concerning caenopore-5 purification.

I would like to express my appreciation to the previous and current members of the lab with whom I shared my time: Dr. Olivier Hecht, Dr. Andreas Schwanter, Dr. Michael Pachta, Dr. Inken Lorenzen, Sascha Jung, Matthias Michalek, Agnes Nolting, Patrycja Kozak, Sonja Hollmer, Britta Hansen, Ahmad Trad, Jonas Völcker, Marc Rödinger and Björn Spudy. They made my working environment inspiring and enjoyable.

Special thanks go to Dr. Michael Pachta, Sascha Jung and Christiane Wiessner for their proof-reading and important remarks to my thesis.

I thank Dr. Radek Sedlacek, Inken Beck and Krzysztof Paliga for their care and support and nice time we spent together.

A very special “thank you” goes to my friends: Lidia Stepien, Monika Gajek, Asia Lewandowska, Aleksandra Fedorska, Dorka Dec, Magda Golda and Kasia Szukalska, who has brought so much fun and so many colors to my life.

Finally I wish to express my deep indebtedness to people who are the most important in my life my mother and my two sisters: Ania and Marysia and brother Janusz for their love and unconditional support.

The Curriculum Vitae

Name:	Justyna Myśliwy
Date Birth:	21.03.1978
Place of Birth:	Jasło (Poland)
Nationality:	Polish

Education:

12/2002 – present	PhD student at the Faculty of Medicine, Department of Biochemistry, Christian Albrechts University Supervisor: Prof. Dr. J. Grötzinger
10/2001 – 06/2002	The experimental part of M.Sc. thesis and M.Sc. degree in molecular biology The title of the M.Sc. thesis: "Analysis of the thiamine pyrophosphokinase from *Saccharomyctes cerevisiae* and *Fagopyrum esculentum*" Supervisor: Prof. Dr. A. Kozik
10/1997 – 06/2001	Undergraduate study: courses in biology, chemistry and molecular biology Faculty of Biology, Jagiellonian University, Cracow, Poland
09/1993 - 06/1997	Student of secondary school (Jasło, Poland)
09/1985 – 06/1993	Student of elementary school (Szebnie, Poland)

Publications:

Mysliwy, J., Dingley, A. J., Sedlacek, R., Grötzinger, J.:
Structural characterization and binding properties of the hemopexin domain of the matrixmetalloproteinase-19. Protein Expres Purif, 2006, 46 (2): p. 406-413.

Declaration

I, Justyna Myśliwy, hereby declare that I produced the present doctoral thesis all by myself; except for scientific advice given by my doctoral thesis supervisor, and that I did not use any other aids besides those that were stated.

Also, this doctoral thesis has not already been submitted to any other faculty.

Furthermore, I declare that I have not unsuccessfully terminated any other doctoral procedure so far and that I have not been deprived of an already acquired degree of a doctor.