The three dimensional structure of plantaricin J

Determination, by NMR spectroscopy, of the three dimensional structure of the plantaricin J component of the plantaricin J/K two-peptide bacteriocin

Mads Christofer Haugen

For the Master Degree at the Department of Molecular Biosciences

UNIVERSITY OF OSLO

November 18. 2008

Acknowledgements

Nuclear magnetic resonance (NMR) spectroscopy was preformed at The Department of Chemistry, University of Oslo, by Dr Per Eugne Kristiansen.

Peptide production and purification was preformed at The Institute for Molecular Biosciences, University of Oslo, under supervision of Dr Gunnar Fimland, Professor Jon Nissen-Meyer and Per Rogne. At the time, they where all part of Professor Jon Nissen-Meyer's research group, Dr Gunnar Fimland is presently working at Rikshospitalet.

NMR sample preparation, spectrum assignment and structure calculation was preformed at The Institute for Molecular Biosciences, University of Oslo, under supervision of Dr Per Eugene Kristiansen and Per Rogne.

I would like to thank all my supervisors for the help they have given throughout the work on this thesis; it has been needed and much appreciated.

I would also like to thank all fellow students and researchers in Professor Jon Nissen-Meyers research group.

Camilla and Leo, I love you!

Abstract

The bacteriocin plantaricin J/K consists of the two peptides PlnJ and PlnK, and the two peptides must be present for optimal bacteriocin function. In this study, the three dimensional structure of the PlnJ peptide was determined by use of NMR spectroscopy. ^{15}N labelled PlnJ for structure determination was produced by *Escherichia coli* BL-21. A pGEV-2 vector, containing a gene encoding a fusion of the immunoglobulin binding domain of Streptococcal protein G (GB1 domain) and PlnJ, was used to transform the bacteria. Peptide production was achieved by inducing expression of the gene in ^{15}N enriched minimal media (M9). The PlnJ target peptide was cleaved from the GB1 domain fusion partner with cyanogen bromide, and purified using reverse phase HPLC. Molecular weight was determined to be 2970, using a MALDI-TOF mass spectrometer, indicating a 99 % degree of ^{15}N labeling. The produced PlnJ was found to be biologically active when mixed with equimolar amounts of synthetic PlnK prior to being exposed to cells sensitive to plantaricin J/K. Structural investigations of PlnJ were made using circular dichroism (CD) spectroscopy and NMR spectroscopy. Structuring was calculated from CD data using the absorbance at 222 nm as a quantitative measure of α-helical content, and by fitting the entire spectrum to the CD spectra of a data base using the CDPro software package. Structural constraints were obtained by analyzing NMR spectra: NOE distance constraints from a NOESY-HSQC spectrum and a NOESY spectrum, dihedral angle constraints calculated from the $^{3}J_{HNH\alpha}$ found in a $H^{N}H^{\alpha}$ spectrum, φ and ψ torsion angle constraints obtained by matching chemical shifts to a data base using the TALOS software. These constraints were used to calculate a three dimensional structure using the CYANA software. All structural data supports an α-helical structure. Furthermore, the data suggests that the helical content is divided between two regions spanning amino acid residues 3-13 and 18-21 and that the molecule is linear. The side chains of the polar and the non polar amino acid residues of PlnJ are confined to separate sides of the long axis of the structure, making it an

amphiphilic molecule.

1. Introduction

1.1 Antimicrobial peptides

Antimicrobial peptides, AMP's, are a divers group of molecules that inhibit the growth of micro organisms. Numerous bacterial strains and species of mammals, amphibians, birds, insects and plants synthesize these substances. AMP's kill micro organisms that are pathogenic to, or compete with, the producing organism and may thus be considered evolutionary adaptations [1-4].

1.2 Bacteriocins

The term bacteriocin first described antimicrobial substances secreted by bacteria [5]. A later paper specified that a bacteriocin had to be of proteinaceous nature and kill bacteria closely related to the bacteriocin producing species [6]. Currently, any ribosomally synthesized AMP from bacteria is termed a bacteriocin [7]

Although in some ways similar to conventional antibiotics, bacteriocins differ from these in that they are ribosomally synthesized. Their target specificity is much narrower, and to each bacteriocin system there is associated a protein conveying immunity. Bacteriocins are produced during the exponential phase of the bacteria's growth while antibiotics are synthesized during the stationary phase [8, 9].

In many of the bacterial species investigated, there have been found bacteriocins [6, 10]. The bacteriocins produced by lactic acid bacteria (LAB) and *E. coli* are the ones that are best characterized, biochemically as well as genetically and structurally. The LAB bacteriocin nisin (*Lactococcus lactis)* has for the past 50 years been used as a food preservative [9]. LAB naturally occur in food consumed by mammals, they are, as their bacteriocins, non pathogenic, and compete with pathogenic and food spoiling

bacteria. The great interest in LAB bacteriocins is largely due to their potential for use as food preservatives and in treatment of disease [9, 11].

1.3 Classification of LAB bacteriocins

LAB bacteriocins may be divided into three classes based on their size, the presence or absence of modified amino acid residues and heat stability [8].

- Class I- Small (<10kD), posttranslationally modified polypeptides that contain the modified amino acid residues lanthionine and/or methyllanthionine. These bacteriocins are often termed lantibiotics.
- Class II- Small (<10kD), polypeptides that do not contain modified amino acid residues.
- Class III- Large (relative to other bacteriocins) (>10kD), heat labile polypeptides.

The unmodified class-II peptide bacteriocins are further divided into three subclasses:

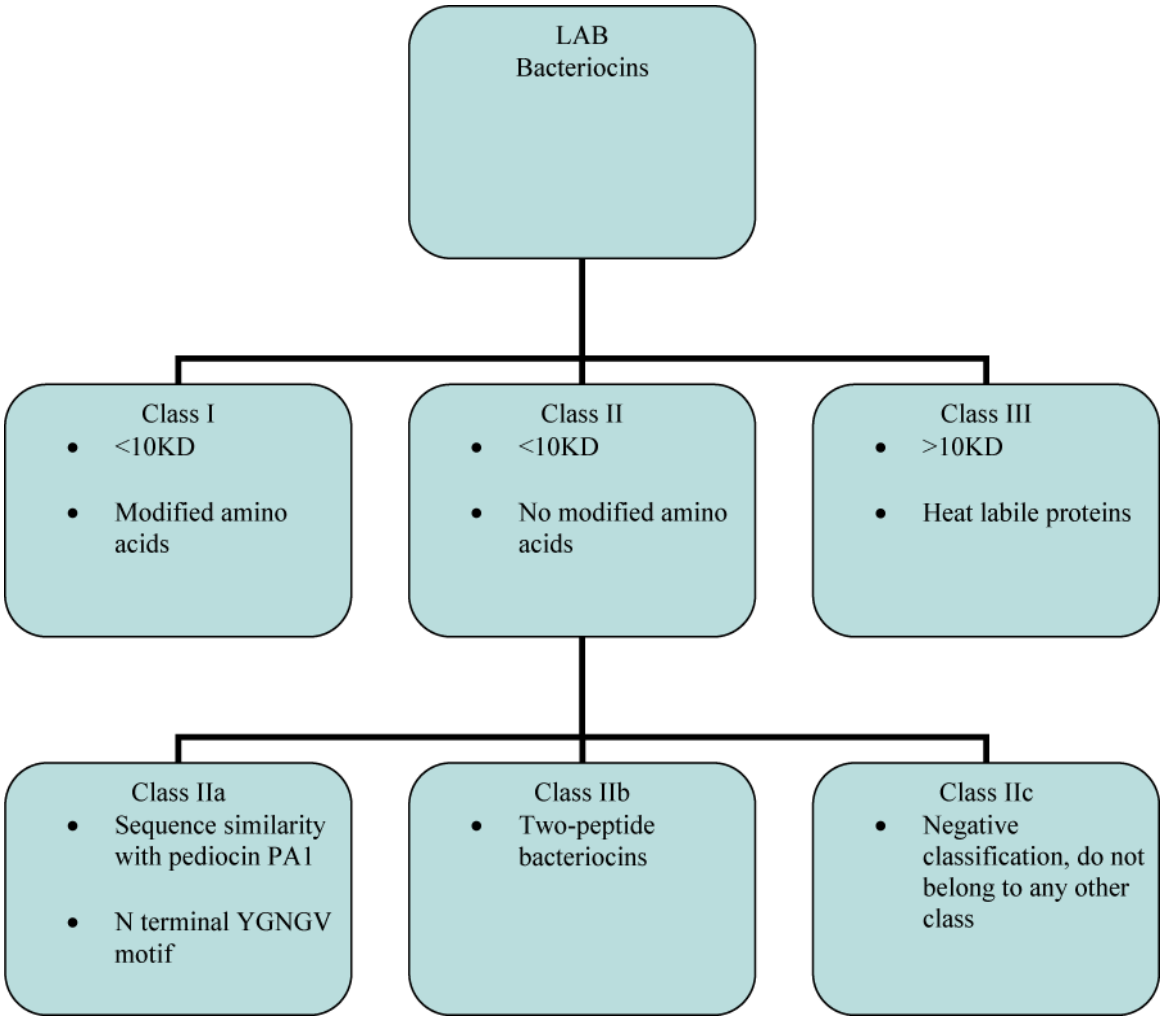

Figure 1 Classification of bacteriocins and sub-classification of the class II bacteriocins.

1.3.1 Class-IIa

The pediocine-like peptides are grouped together because of their sequence similarity, and take their name from pediocin PA-1, the first of the group to be described. All the class-IIa bacteriocins have a conserved YGNGV motif in the N-terminal half [12]. In addition to the conserved motif, the bacteriocins in this subclass show extended sequence similarities, particularly in the N-terminal part, as well as 3D structural similarities. The general structure of the pediocin-like bacteriocins

consists of an N-terminal β-sheet-like structure and a C-terminal helix-containing structure, linked by a hinge region [13].

1.3.2 Class-IIc

Class-IIc consists of a divers group of one-peptide bacteriocins that do not share sequence similarities with pediocin PA-1. Further subdivisions have been suggested for this class [14].

1.3.3 Class-IIb

Class-IIb consists of the two-peptide bacteriocins, all of which contain two synergistically functioning peptides that are cationic at neutral pH and contain hydrophobic and/or amphiphillic regions [15]. Both peptide components of the two-peptide bacteriocins must be present for optimum activity, as the individual peptides have little, if any, bacteriocidal activity [16]. The simultaneous and equimolar secretion of each of the peptides is also important to assure maximum bactericidal effect, and, at optimal conditions, the two-peptide bacteriocins display at least a thousand fold higher activity than each of the peptides separately [15]. As elaborated below, it has been shown that both peptides of all class-IIb bacteriocins are transcribed from genes positioned next to each other within the same transcriptional unit, and that there is only one immunity protein per class-IIb bacteriocin. In addition, it has been documented that two-peptide bacteriocins are in physical contact while performing their antimicrobial function [17, 18]. This indicates that the two peptides of two-peptide bacteriocins function as one antimicrobial unit, and are not simply two synergistically acting one-peptide bacteriocins. Below is a table of two-peptide bacteriocins.

Table 1 Sequences of known two-peptide bacteriocins. GXXXG motif are marked by the bold case. Note that the β peptides of plantaricin S and NC8 lack the GXXXG motif, but have similar AXXXA (pls β) or SXXXS(pln NC8). Sequences from [19]

Bacteriocin	Sequence
Lactococcin G	Lcnα: GTWDDI**GQGIG**RVAYWV**GKAMG**NMSDVNQASRI NRKKKH Lcnβ: KKWGWLAWVDPAYEFIK**GFGKG**AIKEGNKDKWK NI
Lactococcin Q	α: SIWGDI**GQGVG**KAAYWV**GKANG**NMSDVNQASRI NRKKKH β: KKWGNLAWVEPAGEFLK**GFGKG**AIKEGNKDKWK NI
Plantaricin E/F	E: FNRG**GYNFG**KSVRHVVDAI**GSVAG**IRGILKSIR F: VFHAYSARGVRNNYKSAVGPADWVISAVR**GFIHG**
Plantaricin J/K	J: GAWKNFWSSLRK**GFYDG**EAGRAIRR K: RRSRKNGI**GYAIG**YAFGAVERAVLGGSRDYNK

Plantaricin S	α: RNKLAYNM**GHYAG**KATIFGLAAWALLA β: KKKKQSWYAAAGDAIVSFGEGFLNAW
Plantaricin NC8	α: DLTTKLWSSW**GYYLG**KKARWNLKHPYVQF β: SVPTSVYTLGIKILWSAYKHRKTIEKSFFNKGFYH
Lactacin F	A: RNNWQTNV**GGAVG**SAMI**GATVGGTICG**PACAVA GAHYLPILWTGVTAATGGFGKIRK X: NRWGDTVLSAASGAGTGIKACKSFGPWGMAICG VG**GAAIGGYFG**YTHN
Brochocin-C	A: YSSKDCLKDI**GKGIG**A**GTVAGAAGGGLAAG**L**GAIP G**AFV**GAHFGVIGG**SAACI**GGLLG**N B: KINW**GNVGGSCVGGAVIGGALGGLGGAGGGCITG** AIGSIWDQW
Thermophilin 13	A: YSGKDCLKDMGGYALA**GAGSG**ALW**GAPAG**GV**G ALPG**AFV**GAHVGAIAG**GGFACM**GGMIG**NKFM B: QINW**GSVVGHCIGGAIIGGAFSGGAAAG**V**GCLVG**S GKAIINGL

ABP-118	Abp118α: KRGPNCV**GNFLGGLFAG**LAAA**GVPLG**PA**GIVGG**A NL**GMVGG**ALTCL Abp118β: KNGYGGSGNRWVHGGA**GIVGGALIGAIGG**PWSAV A**GGISGG**FTSCR
Salvaricin P	Sln1: KRGPCNV**GNFLGGLFAGAAAGVPLG**PA**GIVGGANL GMVGG**ALTCL Sln2: KNGYGGSGNRWVHCGA**GIVGGALIGAIGG**PWSAVA **GGISGG**FASCH
Mutacin IV	NlmA: KVSGGEAVAAIGICATASAAI**GGLAG**ATLVTPYSVG TWGLIRSH NlmB: DKQAADTFLSAVG**GAASG**FTYCASNGVWHPYILA GCAGVGAVGSVVFPH
Lactocin 705	705α: GSM**GIYQG**IPDFLK**GYLHG**ISAANKHKKGRLGY 705β: **GFWGG**L**GYIAG**RV**GAAYG**HAQASANNHHSPING
Enterocin 1071	A: ESVFSKI**GNAVG**PAAYWILKGLGNMSDVNQADRI NRKKH B: GPGKWLPWLQPAYDFVT**GLAKG**IGKEGNKNKWK NV

1.4 Genetics and expression of class-IIb bacteriocins

In general, five genes are needed for the expression of two-peptide bacteriocins; two structural genes encoding the bacteriocin itself, a gene encoding an immunity protein, a gene encoding a dedicated ABC-transporter and one encoding an accessory protein of unknown function, but which is important for secretion [10, 20, 21]. The structural genes are found next to each other on the same operon, ensuring that equal amounts of the two are produced. The gene for the immunity protein is found on the same operon [16]. All two-peptide bacteriocins are transcribed as inactive preforms containing an N-terminal 15-30 residue double-glycine leader [16]. Upon export, this leader sequence is cleaved off at the C-terminal end of the double-glycine motif, thus activating the bacteriocin [16]. Distinguishing the dedicated bacteriocin ABC-transporter from other ABC-transporters is a 150 residue N-terminal extension [22]. Studies on the lactococcin G two-peptide bacteriocin and its ABC transporter revealed that the N-terminal extension cleaved off the leader sequence of the bacteriocin [22]

1.5 Immunity

Immunity against the bacteriocin is conveyed by a protein transcribed from the same operon as the bacteriocin itself. Putative immunity protein genes have for the most part been confirmed as such upon rendering cells, into which the genes have been transferred, insensitive to their cognate bacteriocin [23-29]. It has not been clarified exactly how the immunity proteins confer resistance to bacteriocins, but studies show that the proteins have to be expressed by the cell, simply adding immunity protein to a cell culture will not deter bacteriocin activity [27]. Even though the class-IIb bacteriocins consist of two peptides, they each have only one immunity protein

1.6 Control of expression

Expression of some of the two-peptide bacteriocins is controlled by a three-component quorum sensing regulatory pathway [30-38]. The components of this pathway are a peptide pheromone, a membrane associated histidine protein kinase and response regulators [39, 40]. Quorum sensing is a cell density sensing system: a low constitutive expression of a peptide pheromone, plnA in the LAB strain *L. plantarum*, leads to increasing pheromone concentration as cell density increases. When the pheromone concentration reaches a threshold value, the histidine protein kinase is triggered to phophorylate two response regulators, which in turn activate genes involved in bacteriocin production [32-37]

1.7 Structure of two-peptide bacteriocins

Circular dichroism (CD) studies combined with Edmundson α-helical wheel sequence displays reveal that a common structural feature of some two-peptide bacteriocins is an amphiphilic α-helix [16]. The structuring of the peptides required the presence of a structure inducing agent, such as tetrafluoroethanol (TFE), or a membrane mimicking entity like dodecylphosphocholine (DPC) micelles. It was also shown that further structuring was achieved upon equimolar and simultaneous addition of the two peptides of two-peptide bacteriocins to the membrane mimicking entities, and it thus appears that the two peptides of two-peptide bacteriocins interact in a structuring manner when exerting their bactericidal effect [16].

Another suggested secondary structure element of two-peptide bacteriocins is a hydrophobic β-sheet either with or without a cystein residue at each end of the primary structure [41].

A reoccurring sequence motif in many, if not all, two-peptide bacteriocins is a GXXXG (AXXXA and SXXXS also occur) sequence within a stretch that may form an α-helix, see *Table 1*. The additional structuring seen in two-peptide bacteriocins upon interaction of their two peptides with membranes, is thought to be mediated by the GXXXG motifs since it has been shown that these motifs are involved in helix-helix interactions [42]. Mutational studies have shown that substituting the glycines in some of these motifs with more bulky amino acid residues is detrimental to the activity of the two-peptide bacteriocins enterocin 1071 and lactococcin G [43], indicating that the glycines facilitate close contact.

For lactococcin G, a structure has been proposed in which the α and β peptides are connected through the GXXXG motifs of residues 7-11 (α) and 18-22 (β), and the two peptides form a coil-coil and lay in a parallel yet staggered way relative to each other [44]. In this proposed structure, the GXXXG motif of the α-peptide lies in an α-helix spanning residues 3-21, whereas the β-peptide GXXXG motif lies at the end of an α-helix spanning residues 6-19 [44]. There are C-terminal helixes in both peptides, residues 24-34 in the α-peptide and residues 23-32 in the β-peptide. The C-terminal helix of the β-peptide was only observed when TFE was used as structuring agent, and the above proposed GXXXG interacting structure could cause structuring of residues 18-21 and stabilization of the C-terminal helix of the β-peptide. This would lead to a continuous helix-helix structure running from the N-terminus to residue 24 in the α-peptide, and from residue 12-35 in the β-peptide [44]. It is further suggested that this might be a general structure for a group of three homologous two-peptide bacteriocins; lactococcin G, lactococcin Q and enterocin 1071 [44]. The suggested structure accounts for the enhanced helix formation upon α-β interaction reported by Hauge et.al. [17].

NMR studies of the plantaricin E/F two-peptide bacteriocin reveal that its two peptides both adopt amphiphilic α-helical structures in the presence of DPC micelles. In PlnE, there is a kink at proline 20, whereas for PlnF the helical structure is

interrupted by a more flexible region at amino acid residues 21-24. In both peptides the N-terminal region is both more polar and less structured [45].

1.8 Mode of action

All two-peptide bacteriocins, whose mode of action has been analyzed, render the membrane of target cells permeable to small monovalent molecules and thus disturb vital transmembrane concentration differences. Insight as to how ion conducting pores are made is still lacking, but the fact that membrane-mimicking substances induce structuring in bacteriocins indicates a membrane-interacting mode of action [16].

Of the two-peptide bacteriocins who's mode of action has been investigated, there has been shown to be a difference in ion specificity of the pores formed in the membrane of target cells [16]. This specificity, combined with additional structuring of the bacteriocins upon membrane interaction and the low peptide concentration needed for activity, leads to the conclusion that a general destabilization of target cell membranes as a result of peptide interaction is not plausible [16].

A membrane spanning target cell interaction has been suggested for lactococcin G: a cluster of positively charged amino acid residues at the C-terminus is thought to be driven through the membrane by the transmembrane potential (negative on the inside), leaving the double helix inside the membrane and the N-terminus outside or in the inter phase of the membrane [44].

1.9 Plantaricin J/K

Table 2 Amino acid sequence of plantaricin J/K, one-letter code. GXXXG motif in bold.

PlnJ GAWKNFWSSLRK**GFYDG**EAGRAIRR
PlnK RRSRKNGI**GYAIGYAFG**AVERAVLGGSRDYNK

Table 2 shows the amino acid sequence of plantaricin J/K.

The molecular mass of plantaricin J (PlnJ) is 2929 Da, and that of plantaricin K (PlnK) is 3503 Da.

Plantaricin J/K is one of several peptide bacteriocins produced by *Lactobacillus plantarum C11*, all of them referred to as plantaricins. Investigations show that there seems to be a total of five operons involved in the production of mature bacteriocin in *L. plantarum C11*: *plnABCD*, *plnEFI*, *plnJKLR*, *plnMNOP* and *plnGHSTUV* [31]. The secreted peptide product, Pln A, of the A gene on the ABCD operon is the peptide pheromone that induces both transcription of the ABCD locus and bacteriocin production [32]. The *plnB* gene encodes a histidine kinase, which together with two antagonistic response regulators, encoded by the *plnC* and *plnD* genes, effect the transcription of genes induced by PlnA [34]. Of the remaining four operons, three contain ORF's for bacteriocins, PlnEF, PlnJK and PlnN. PlnEF and PlnJK are two-peptide bacteriocins, whereas PlnN is a one-peptide bacteriocin. Included on the operon with each of these bacteriocins are the genes encoding their corresponding immunity proteins: PlnI, PlnL and PlnM/P respectively [31]. The forth operon, *plnGHSTUV*, contains the GH ORF encoding an ABC transporter, PlnG, and its accessory protein, PlnH [31].

1.10 Aim of study

The aim of this study is to determine the three dimensional structure of the plantaricin J peptide-component of the plantaricin J/K two-peptide bacteriocin by NMR spectroscopy. In order to accommodate ^{15}N edited NMR techniques, a ^{15}N marked peptide will be produced by over-expression in *E. coli* using ^{15}N enriched media. Before NMR spectroscopic investigation of the peptide, it will be purified using HPLC, molecular weight will be determined by MALDI-TOF mass spectroscopy, and biological activity will be verified through a bacteriocin assay.

Structure determination is the key to understanding how a bacteriocin exerts its function. Primary structure, secondary structure and the orientation of secondary structure elements relative to one another provide a topographic map of the peptide, which is important in understanding how the bacteriocin interacts with the target cell membrane when forming ion specific pores.

As mentioned in section 1.7, it has been shown that the two peptides of two-peptide bacteriocins are in contact when exerting their bacteriocin activity and that further structuring is achieved when the two peptides make contact. The determinations of the individual structures are therefore necessary steps to elucidate the structure-function relationship of plantaricin J/K.

Bacteriocins are unstructured in water, but gain structure as they make contact with the membrane of the target cell. In order to induce structuring of the peptide, DPC micelles serve as a membrane mimicking entities [46].

2. A brief description of some methods used in this thesis

2.1 Circular Dichroism (CD) spectroscopy

Due to the different extinction coefficients optically active samples have for left and right circularly polarized light, absorption produces elliptically polarized light. A circular dichroism spectrum is a display of either the per residue difference in extinction coefficient, for right and left circularly polarized light, Δε, or the per residue elipticity, θ, as a function of wavelength [47].

θ is defined as

$$\tan\theta = \frac{E_R - E_L}{E_R + E_L}$$

Formula 1

where E_R and E_L are the electric field vector magnitudes of right and left circularly polarized light, respectively [47]. This is further illustrated by the figure below:

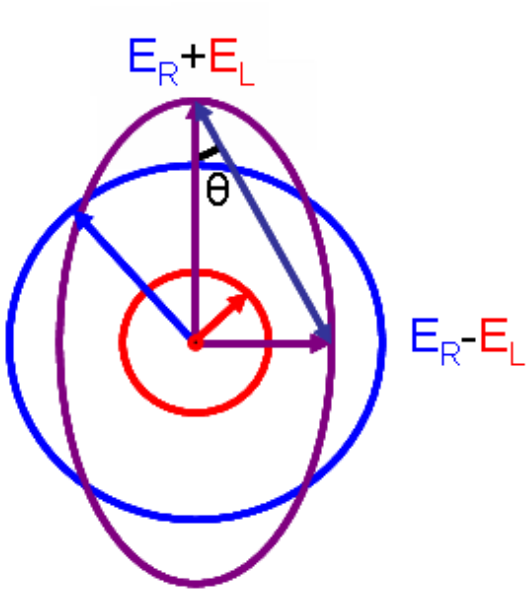

Figure 2 Unequal contributions of right (red)- and left(blue)- circularly polarized light creates elliptically polarized light of θ degrees. Figure from Wikipedia.

There is a simple proportional relationship between these sizes as they both can be used to calculate the difference in absorption of right and left polarized light, ΔA;

$$\Delta\varepsilon cl = \Delta A = \theta / 32982$$

Formula 2

where c is concentration and l is the length of the light path. Chiral centers in amino acids give proteins an intrinsic optical activity, but more interesting is the fact that the asymmetry of the α-helix contributes to this activity, thus making it possible to calculate helical content by CD spectroscopy [47].

Figure 3 below shows the signature CD spectra of an α-helix, a β-sheet and a random coil protein.

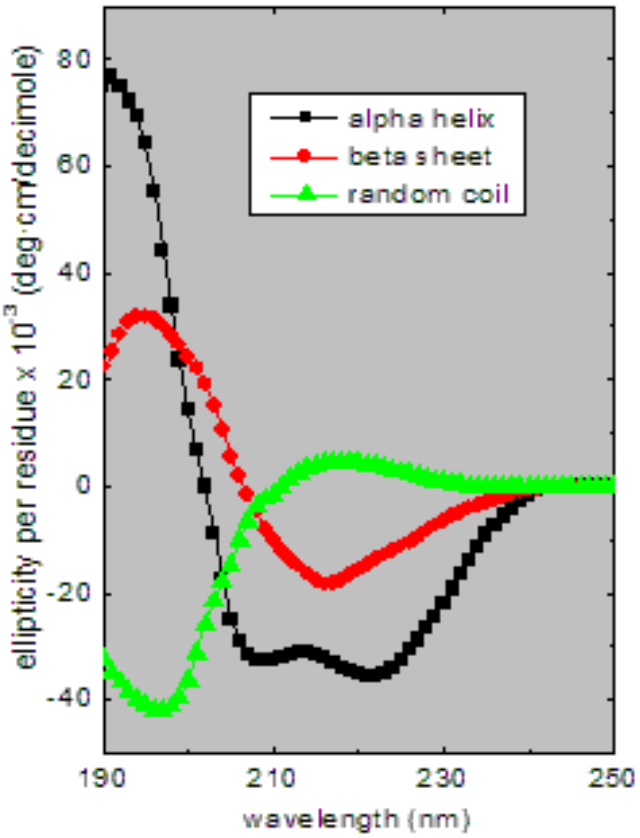

Figure 3 Signature CD spectra of α-helix, black, β-sheet, red, and random coil protein, green. Figure from www.ap-lab.com/circular_dichroism.htm

2.2 Nuclear Magnetic Resonance spectroscopy

2.2.1 Basics

NMR active nuclei possess a property called spin angular momentum, **I**. It is characterized by the spin angular momentum quantum number, I, called the spin of the nucleus, which can take integer or half integer values ≥ 0 [48].

This thesis will only consider spin ½ nuclei.

In a Cartesian coordinate system, the z component of **I**, I_z, has a value specified by the magnetic quantum number, m:

$$I_z = \hbar m$$

Formula 3

where ħ is Plancks constant divided by 2π [48].

For spin ½ nuclei there are two possible states for I_z; 1/2ħ and -1/2ħ, these are referred to as the α and β state, respectively [48].

Aligned with **I** (when **I** > 0) is a nuclear magnetic moment, **μ**. The z component of **μ**, μ_z, has a value which is a function of I_z and the magnetogyric ratio, γ, intrinsic to each nucleus [48]:

$$\mu_z = I_z \gamma$$

Formula 4

The quantization of I_z leads to quantization of μ_z [48].

When a system of spins is placed in a static magnetic field, B_0, the magnetic nature of μ_z leads to non-degeneracy of the two I_z states α and β. The energy difference between the α and β states is proportional to both the static magnetic field and γ. The sign of gamma determines which of the states is the low energy one [48].

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

Formula 5

The energy difference leads to a difference in the population of the two states. The fraction of the population of the two states, Nα/Nβ, is described by the Boltzman distribution:

$$N\alpha / N\beta = e^{-\Delta E / kT}$$

Formula 6

where k is the Boltzman constant and T is the absolute temperature [48]. The difference in population leads to an accumulation of the z-components of the magnetic moments along or against (positive and negative γ respectively) the static magnetic field. Since the value of I_z is less than the value of I, it follows that **μ** is not aligned with B_0, but the x and y components of **μ** are non-coherent and sum to zero [48]. This results in a net magnetization vector **M**, aligned with the static magnetic field [48]. The rotating frame of reference, a Cartesian coordinate system rotating with the same frequency as the Larmour frequency of the observed spin is defined as having its +z-axis pointing in the same direction as **M.** *Figure 4* illustrates the net magnetization vector.

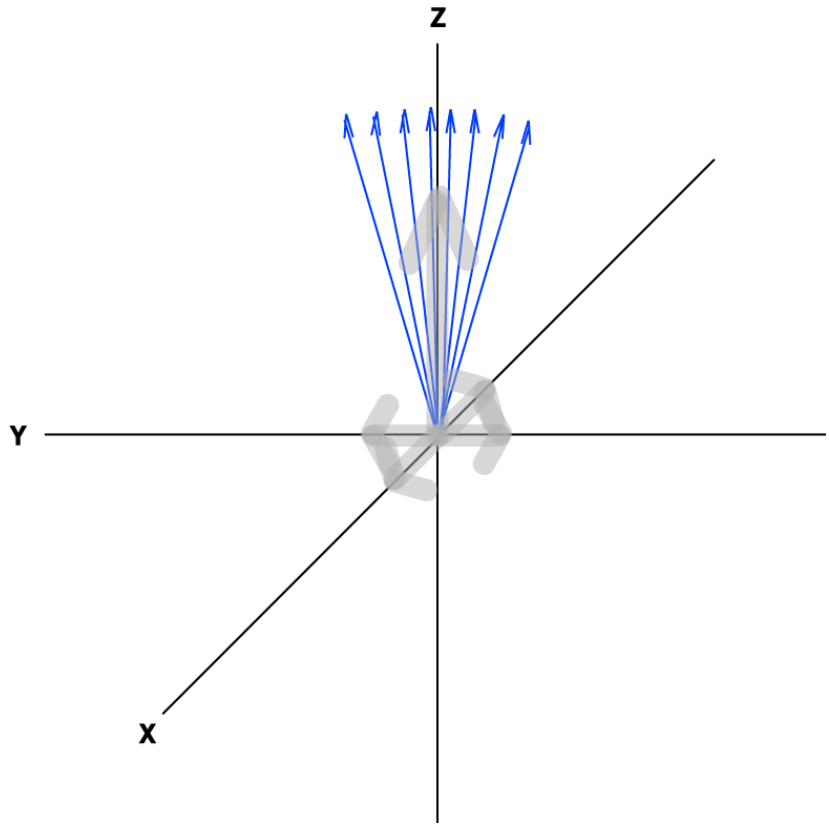

Figure 4 Illustration of net magnetization vector, M. Blue arrows are individual spin magnetization vectors, each having x-,y- and z-components. Grey arrows are net magnetization vectors for all axes. All net magnetization vectors, except the one in +z direction, have counterparts pointing in the opposite direction. This means that the effect of all magnetization, except +z magnetization, is canceled by magnetization pointing in the opposite direction.

The static magnetic field imposes a torque on **μ**, causing it to precess about the direction of the static magnetic field, which coincides with the z-axis [48].

The rate with which **μ** precesses is proportional to the energy difference between the two I_z states [48]:

$$\Delta E = \hbar \nu$$

Formula 7

$$\nu = \gamma B_0 / 2\pi$$

Formula 8

This is the Larmour frequency of the spin.

The γ-values describe the magnetic strength of the nuclei [49]. Below is a table of important NMR active nuclei, their natural abundance and γ values.

Table 3 The magnetogyric ratio and natural abundance of some interesting spin 1/2 nuclei

Nuclei	Magnetogyric ratio $10^7 radT^{-1}s^{-1}$	Natural abundance %
^{1}H	26.75	99.98
^{13}C	6.728	9.11
^{15}N	-2.713	0.37
^{31}P	10.83	100
^{19}F	25.18	100

In the following, a spin refers to the magnetic moment associated with spin ½ nuclei.

2.2.2 Chemical shift

Local secondary magnetic fields, arising from motion in the local electron density induced by the static magnetic field, alters the perceived static field, B_{local}, for all spins. It is said that local magnetic fields either shield or deshield spins from the effect of the static magnetic field. For isotropically tumbling molecules, this shielding/deshielding is described by adding a modifying term to the function for the Larmor precession of spins. The modifying term contains σ, which is an average isotropic shielding constant for each spin [48].

$$\nu = -\gamma / 2\pi(1-\sigma)B_0$$

Formula 9

Positive σ describes a shielding effect in which the spin in question perceives a weaker magnetic field and hence precesses at a lower frequency. The shielding is caused by local magnetic fields opposing the static field [48]. This shift in resonant frequency due to a chemical environment is called the chemical shift of spins.

In proteins/peptides, the resonant frequency of spins varies in a predictable manner depending on their position in an amino acid and the structure of the protein/peptide (see 2.2.4) [50-52].

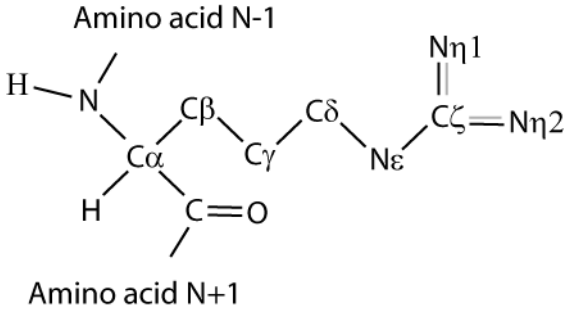

Figure 5 The amino acid arginine in a peptide bond with amino acids N-1 and N+1. The atoms of the side chain are named with Greek letters according to their distance from the backbone. Not apparent from this figure is the geometry about the central C^{α}, see 3.8.5. Nomenclature in the figure deviates from the IUPAC recommendation to give atom names as superscripts to the atom type, this is done to make the figure easier to read.

The atoms of the side chain are given Greek letter names, starting from the C^{α} in *Figure 5*, to identify their position in the side chain. IUPAC suggested names are given as superscripts to the atom type [53]. Numbers are added to differentiate between atoms that are equally distant from the C^{α}, like the two amine nitrogen of arginine, $N^{\eta 1}$ and $N^{\eta 2}$. The protons on $N^{\eta 1}$ and $N^{\eta 2}$ are given a number to reflect which nitrogen they are attached to, and a number to distinguish them from each other. The protons on $N^{\eta 1}$ are thus $H^{\eta 11}$ and $H^{\eta 12}$, whereas the protons on $N^{\eta 2}$ are $H^{\eta 21}$ and $H^{\eta 22}$. The atoms involved in the peptide bond are given names according to the

functional group they belong to. Thus, the two protons (H) in the figure are the amid proton, H^N and the H^α, the carbons are the carbonyl carbon, designated C, and the carbons of the side chain, C^α and on, the nitrogen are the amid one, designated N, and the nitrogen of the side chain, N^ε and on, the oxygen is the carbonyl one, designated O'. The CYANA software utilized in the analysis of NMR data in this thesis uses a slightly different nomenclature: there are no Greek letters and no superscript. Consequently, the H^α is called HA, H^N is H. Atoms with equal distances to CA (C^α) are given numbers like in the IUPAC nomenclature; the protons on NH1 ($N^{\eta 1}$) and NH2 ($N^{\eta 2}$) are thus called HH11 ($H^{\eta 11}$), HH12 ($H^{\eta 12}$), HH21 ($H^{\eta 21}$) and HH22 ($H^{\eta 22}$). In cases where there are two or more equivalent protons, like the three HB (H^β) of alanine, and one can not determine which one gives rise to what resonance, or they all have the same chemical shift, H is substituted with Q. In the text, IUPAC nomenclature will be used, but when referring to CYANA output, CYANA nomenclature will be used with IUPAC nomenclature in parenthesis.

2.2.3 Chemical shift referencing

The dependency of the Larmor frequency on the static magnetic field, leads to a dependency of the chemical shift on the static magnetic field [49]. The use of a common internal standard allows chemical shift referencing independent of field strength, and for this reason Wishart and colleagues compared different compounds in current use, and found sodium 2, 2-dimethyl-2-silapentan-5-sulphonate (DSS) to be a good referencing compound [54]. Chemical shifts are reported as an offset, δ or ppm, from the referencing compound frequency:

$$\delta(ppm) = (\sigma_{reference} - \sigma)10^6$$

Formula 10

where σ is the resonant frequency of the spin in question, Ω, divided by the transmitter frequency ω [48].

2.2.4 J-coupling

J-couplings are through bond interactions between spins, mediated by the electrons that make up the bond. In general, all spins involved in some sort of bond, covalent or hydrogen, are J-coupled. The magnitude of the coupling is described by the J-coupling constant, $^{n}J_{a\text{-}b}$, where n denotes the number of bonds between spin a and b [48].

The coupling modifies the energy levels of the coupled spins; consequently modifying the resonant frequencies of the spins, see *Formula 7*. Two spins, a and b, give rise to two resonant frequencies if uncoupled, the same spins give rise to four resonant frequencies if coupled. This is called a splitting of spins by J-coupling. The split resonant frequencies are centered, in pairs, around the uncoupled a and b frequencies, deviating from these with $\frac{1}{2}\,^{n}J_{a\text{-}b}$ in positive and negative directions, see *Figure 6* [48].

The magnitude of a ^{3}J-coupling is dependent upon the dihedral angle formed by the three bonds. This leads to the possibility of gaining information about protein/peptide dihedral angles, in particular the dihedral angles of the backbone, from coupling constants [48]. See section 3.8.4 for an explanation of backbone dihedral angles.

Below is a schematic presentation of the energy states of a spin.

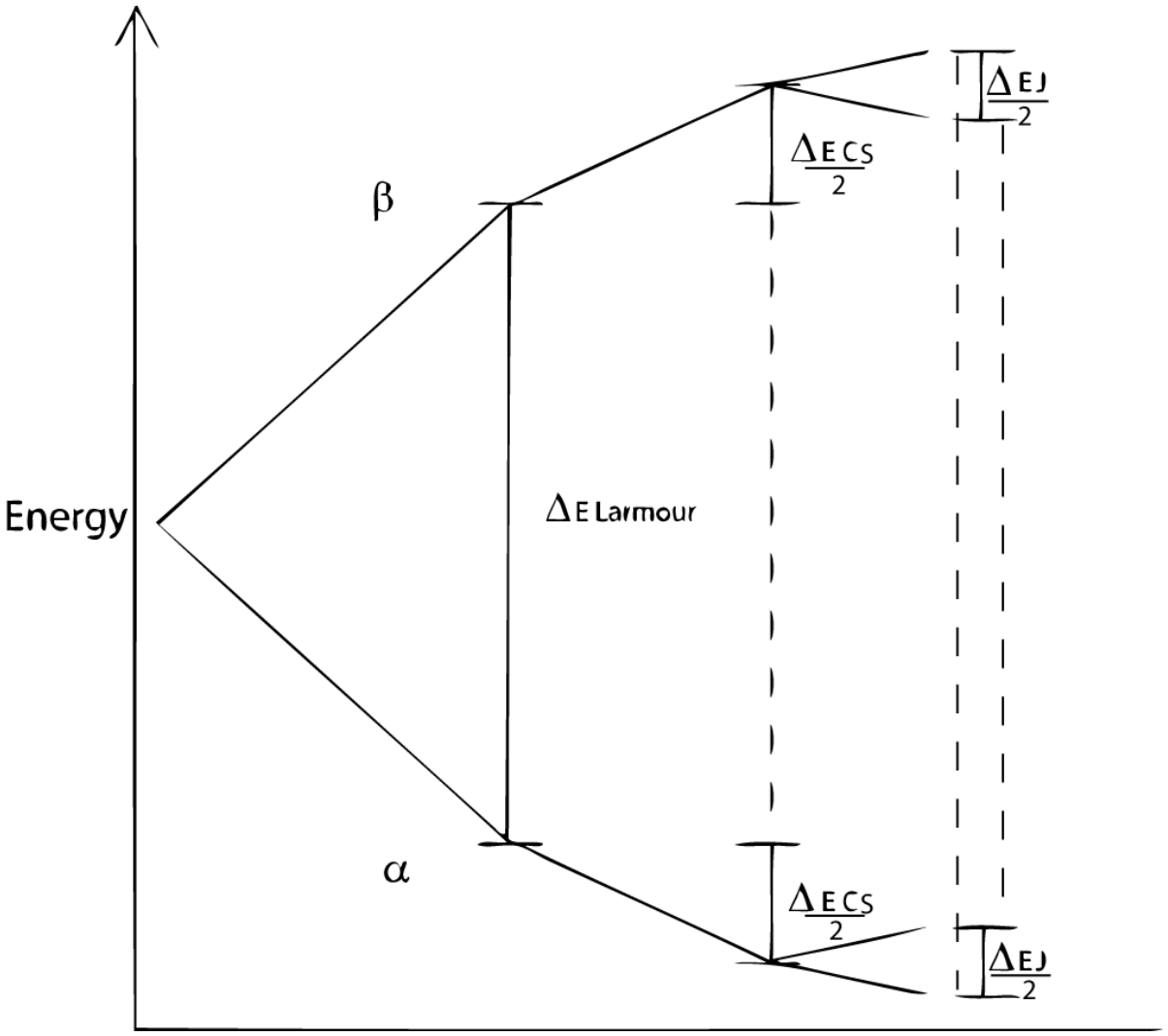

Figure 6 The static magnetic field leads to non-degeneracy of the α and β energy states of a spin. The chemical environment of the spin will shield or deshield the spin from the effect of the static magnetic field causing a decrease (not shown in figure) or increase in ΔE Larmour by ΔE CS (Chemical Shift). A J-coupling will further modify the chemically changed Larmour frequency of a spin by adding and subtracting ¼ Δ EJ from the high and low energy states of the spin. The two allowed transitions between the new energy states, the dashed lines, differ from the chemically shifted ΔE Larmour by ½ ΔE J in positive and negative direction. The energy difference is proportional to the resonant frequency of the spin.

2.2.5 Relaxation

Once perturbed from their equilibrium state Boltzmann distribution, the spins will return towards this equilibrium state through a process known as relaxation [48]. In general, relaxation is described as two steps reversing the effect of perturbation by a

pulse; the coherence gained is lost through what is known as spin-spin relaxation, and the distribution across the energy states is restored through spin-lattice relaxation [48]. Spin-spin relaxation results in loss of detectable signals and is generally the faster of the two, whereas consecutive scans cannot be preformed until the spin-lattice relaxation is complete, restoring +z-direction magnetization [48].

Relaxation affects the line width and sensitivity of spectra, and dictates the minimum time between consecutive scans and must thus be taken into consideration when designing an experiment, but the effects of relaxation on the finished spectrum also offers information about the relaxing molecule [48], see section 2.2.6 about the Nuclear Overhauser Effect below.

2.2.6 Nuclear Overhauser Effect, NOE

The NOE is a change in a spin's signal strength caused by cross-relaxation of dipolar coupled spins [48].

A dipolar coupling is a through space effect spins have on each other due to their dipolar magnetic moments [48].

Cross relaxation is the dependency of the temporal evolution of one spin's longitudinal magnetization on the state of another spin's longitudinal magnetization [48].

In theory, the intensity of the NOEs are inversely proportional with the sixth power of the distance between two nuclei. There are, however, effects that cause the NOEs to deviate from this distance proportionality. The varying chemical environments of nuclei in a molecule will in turn cause varying relaxation pathways; some nuclei will rely heavily on dipol-dipol relaxation, increasing NOE intensities, whereas others might relax through alternative mechanisms, decreasing NOE intensities. Differences in internal molecular motions may also lead to differences in NOE growth rates.

Dipol-dipol interactions tend to be dominated by the pair of nuclei that are closest to each other, so a NOE from one nuclei to a second may be reduced by a third that is closer. Internal distances may also vary over time as molecules assume different conformations [49].

It is therefore with caution NOEs are used to give specific intra molecular distances. NOE data are usually used to give upper distance limits for the separation of nuclei.

2.2.7 Data acquisition and prossesing

An electromagnetic wave with a direction of propagation perpendicular to the above defined z-axis will create a magnetic field perpendicular to that axis. This will cause the spins to start a precession about the direction of this new magnetic field [49]. In NMR terms, this is a pulse. The power, duration and phase of the pulse dictates what direction in the coordinate system the net magnetization will have as a result of the pulse [49]. The frequency of the pulse dictates which type of spin it affects, as the frequency has to match the precessional frequency of that nuclei [49].

The detector of an NMR instrument is made such that it can only detect magnetization in the x-y plane of the above defined coordinate-system [49]. At equilibrium there is magnetization in the x-y plane, but magnetization vectors pointing in all directions cancel each other out, leaving no detectable net magnetization [49]. Pulses, or sequences of pulses with or without time gaps between them, are used to perturb magnetization from equilibrium to a desired state and to transfer magnetization to the x-y plane. Magnetization transferred to the x-y plane has phase coherence as a consequence of the pulses, this leads to detectable net magnetization in the x-y plane [49].

The Free Induction Decay (FID) is a time dependent display of an oscillating voltage in the receiver coil of the detector, caused by precessing magnetization in the x-y plane [49]. This time dependent signal has to be transformed into a frequency domain signal. This is done via Fourier transformation of the digitized signal. *Figure 7* gives a schematic presentation of the effect of the Fourier transform.

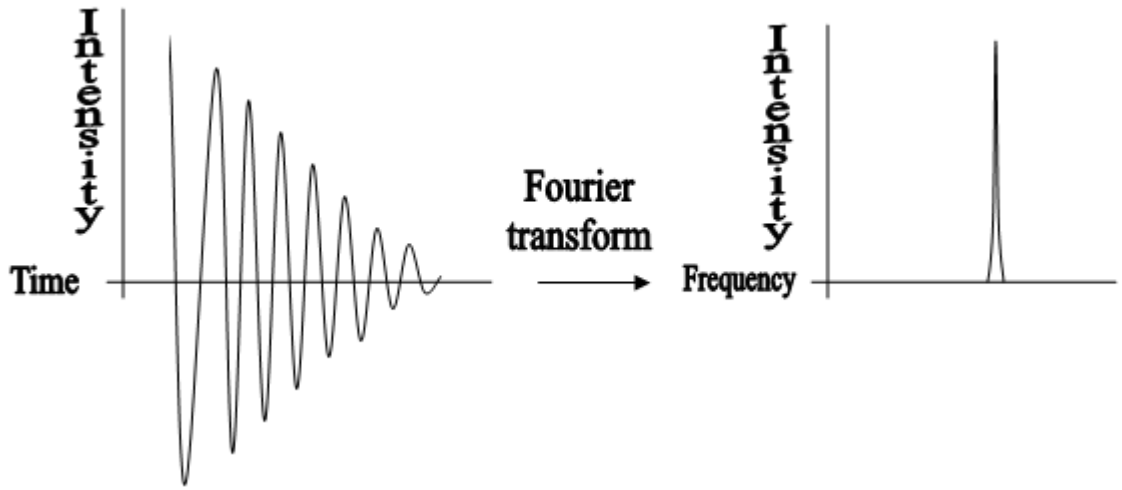

Figure 7 The time dependent Free Induction Decay is Fourier transformed to display the intensities of the signal as a function of frequency.

The Fourier transform can be preformed by a number of algorithms. It is based on a general relationship between a time domain function, s(t), and a frequency domain function S(ν) [48]:

$$S(\nu) = \int_{-\infty}^{\infty} s(t) e^{-i2\pi\nu t} dt$$

Formula 11

2.2.8 Multi Dimensional NMR

The unprocessed NMR signal, the FID, is a time domain display of oscillating currents caused by nuclei precessing in the detector coil. To obtain net magnetization to detect in the x-y plane, nuclei are perturbed from equilibrium by an electromagnetic pulse. The fourier transformed FID gives rise to a one dimensional NMR spectrum [49]. To obtain greater dimensionality, one has to perturb the nuclei an additional time and introduce a variable time (t_1) block between the pulse blocks. *Figure 8* gives a schematic presentation of a two-dimensional NMR experiment. An explanation of the different blocks is given beneath the figure.

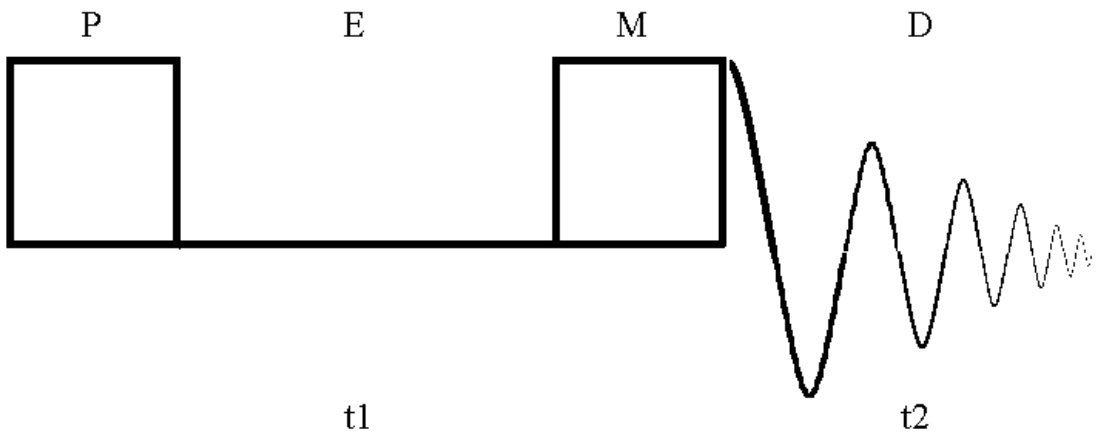

Figure 8 Schematic display of a 2D NMR experiment. Between the preparation pulse, P, and the mixing pulse, M, there is an evolution block, E. The duration of the E-block, t_1, is altered between consecutive scans. Thus a second time dependency is introduced into the FID, giving rise to a second dimension.

The preparation pulse-sequence (P) manipulates the system, ensuring a desired state of magnetization which is to evolve (E) during the t_1 block. The specific sequence may be a single pulse or a complex array of pulses, depending on the type of experiment[49].
The mixing pulse-sequence (M) transfers coherence between spins. The specifics of the sequence determine through which couplings coherence is transferred, and thus which correlations are seen in the finished spectrum [49].

During the detection block (D), the magnetization is measured as a function of time (t_2). As t_1 is augmented throughout consecutive scans, information about the effect of these increments is gathered, i.e. magnetization is measured as a function of time. Double Fourier-transformation of these two time-domains, t_1 and t_2, gives two frequency-domains, a 2D spectrum [49].

The spins are said to be F_1 frequency labeled during the t_1 period. This means that the frequencies the spins precess with during t_1 are the ones displayed in the F_1 dimension of the final spectrum [48].

Additional variable time blocks can be added followed by additional mixing in order to increase the dimensionality further [48].

2.2.9 TOtal Correlation SpectroscopY, TOCSY

TOCSY is a 2D spectrum with proton frequencies in both dimensions used to map continuous chains of J-couplings, e.g. the J-coupling chains in amino acid side chains.

After a 90^o pulse, a spin-lock pulse sequence stops all chemical shift evolution. During the spin-lock, correlation is transferred through J-couplings, but, as all nuclei have the same chemical shifts, they are indistinguishable and the J-coupling is relayed until there is a break in the continuous chain of coupled nuclei or the spin lock is turned of. Approximately 15 ms is needed to relay J-coupling between protons on neighboring carbons [49].

The pattern of peaks from protons J-coupled to H^N is called the fingerprint of the amino acid, TOCSY is used to display the entire fingerprint in one spectrum. However, TOCSY cross peak intensities vary, as the mixing time (spin-lock sequence) varies. Therefore, a series of spectra with increasing mixing time is often used. This is done both to make sure all peaks are found, and to monitor cross-peak intensities as a function of mixing time. Protons distal (proximal) to H^N may have higher (lower) intensities at longer mixing times. This means that cross-peak intensities relative to mixing time may give information about which protons the different peaks originate from [48].

2.2.10 Hetronuclear Single Quantum Correlation, HSQC

HSQC is a 2D spectrum with proton frequencies in the F_2 dimension and the frequencies of heteronuclei in the F_1 dimension [48].

An Insensitive Nuclei Enhancement through Polarization Transfer (INEPT) sequence transfers proton magnetization to the nuclei to which the proton is bound (x-nuclei). A reverse INEPT sequence transfers magnetization back from the x-nuclei to the protons for detection after evolution [48].

HSQC is used to map ^{1}J-couplings between protons and the x-nuclei they are attached to, x being for example ^{13}C or ^{15}N. HSQC is also used to expand 2D proton-proton experiments to 3 or 4D, and thus to separate overlapping proton peaks based on their more highly resolved heteronuclear coupled partners [48].

2.2.11 Nuclear Overhauser Enhancement SpectroscopY, NOESY

NOESY is a 2D proton-proton spectroscopic method. A sequence of two pulses moving the net magnetization 90^{o} in the coordinate system, separated by an evolution delay, t_1, generates F_1 frequency labeled spins and restores magnetization in the +z direction. The spins then evolve under the influence of the dipolar couplings for a duration known as the mixing time of the experiment, before a final pulse creates detectable magnetization [48].

The NOESY spectrum consists of diagonal and off-diagonal peaks. The diagonal peaks reflect the chemical shift of the spins, whereas the off-diagonal peaks correlate dipolar coupled spins [48]. The off-diagonal peaks thus contain information about inter-molecular distances, see section 2.2.6.

2.2.12 HNHA

HNHA is a 3D spectrum correlating H^N, H^{α} and ^{15}N of one residue. The ratio of the intensities of the diagonal H^N-H^N-^{15}N peak and the H^{α}- H^N-^{15}N crosspeak provides a measure for the $^{3}J_{HNH\alpha}$ coupling constant:

$$^{3}J_{H^{N}H^{\alpha}} = \arctan\frac{\sqrt{-\frac{S_{cross}}{S_{diagonal}}}}{2\pi\zeta}$$

Formula 12

S_{cross} and $S_{diagonal}$ are the integrated volumes of the H^{α}- H^{N}-^{15}N crosspeak and the H^{N}-H^{N}-^{15}N diagonal peak, respectively. ζ is the evolution delay, a pulse-sequence parameter [55].

The dihedral angle between H^{N} and H^{α} is found by fitting the coupling constant to a Karplus curve:

$$J_{H^{N}H^{\alpha}} = A + B\cos\theta + C\cos^{2}\theta$$

Formula 13

where θ is the dihedral angle [55].

2.2.13 3D, TOCSY-/NOESY- HSQC

In order to increase resolution, a $^{15}N/^{13}C$ HSQC pulse sequence is incorporated into these two 2D proton-proton techniques. As the low natural abundance of ^{15}N and ^{13}C and the low γ of ^{13}C rises signal strength issues, these techniques are usually preformed on isotope enriched samples [48]. The resulting spectrum is like a stack of proton-proton spectra, one per $^{15}N/^{13}C$ chemical shift. I.e. protons coupled to heteronuclei with different chemical shift will be in different layers of the stack [48]. Below is an illustration of the resolution gains with this method:

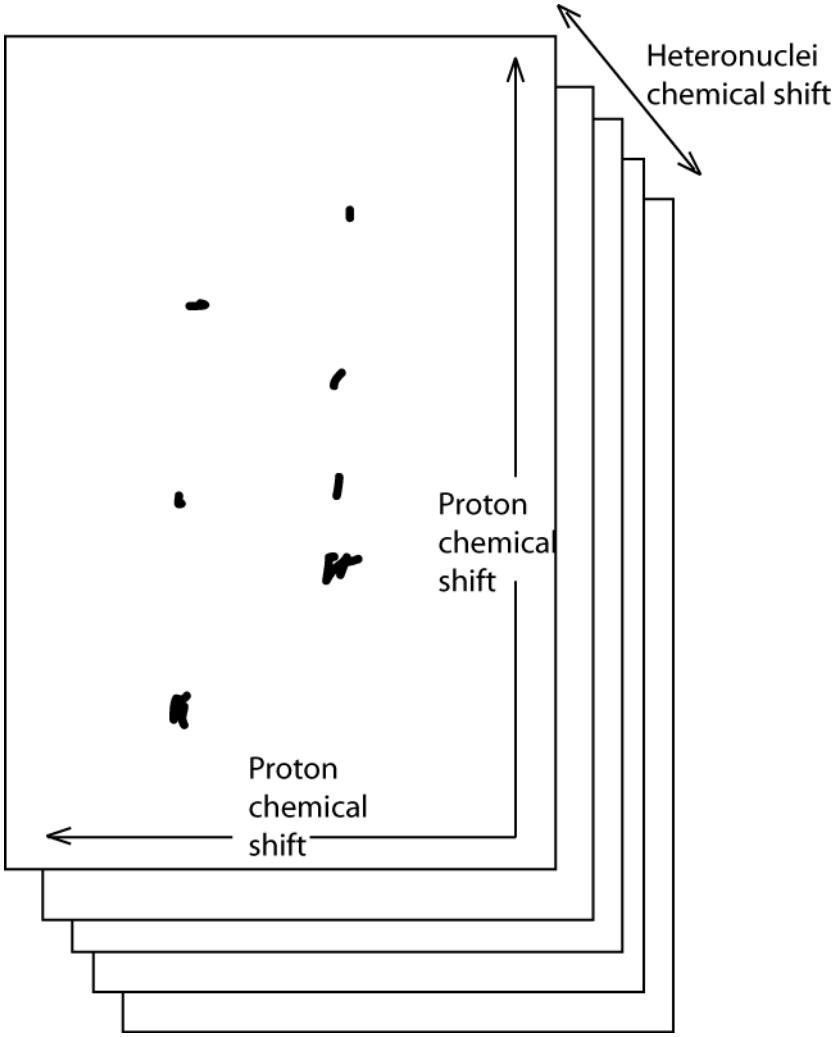

Figure 9 Although the heteronuclei chemical shifts are continuous, the resolution gains of a 3D NOESY-/TOCSY- HSQC spectrum is often illustrated as having a stack of 2D spectra where each layer represents a specific heteronuclei chemical shift.

3. Materials and methods

3.1 Vector DNA

The expression vector used in the work for this thesis was pGEV-2, into which was inserted a gene fusion encoding a fusion protein consisting of, N-terminaly, the immunoglobulin binding domain of Streptococcal protein G (GB1 domain), and, C-terminaly, PlnJ. The vector with gene inserts as described was kindly provided by Per Rogne (PhD student, IMBV, University of Oslo, Norway). The target peptide-GB1 fusion expression system has been specifically designed for fast, high yield expression of small peptides for NMR characterization [56]. The pGEV-2 vector, with out gene inserts, was provided by Dr Gronenborn (National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, USA). The vector contains a strong inducible promotor with a downstream gene encoding the GB1 domain and a linker region. The linker region contains a protease factor Xa cleveage site. However, this cleavage site was not used to cleave off mature PlnJ. Instead, the vector was modified such that the codon for methionine, ATG, was placed immediately upstream of the first codon of PlnJ [44], thus facilitating the use of CNBr to cleave the fusion peptide.

3.2 Preparation of competent *Escherichia coli* cells, $CaCl_2$ method

0.5 ml over night *E. coli* culture was added to 25 ml fresh LB media (see Appendix) and incubated at 37^{o}C until the optical density at 600 nm ($O.D_{600}$) (see Appendix) reached 0.3 (*E. coli* BL-21) or 0.5 (*E. coli* TB-1). The culture was cooled on ice, 10 minutes, and the cells were pelleted (centrifugation at 5000 rpm and 4^{o}C for 10 min). Cells were then resuspended in 5 ml ice cold 0.1 M $CaCl_2$ solution and kept on ice for 5 minutes before they were pelleted again (centrifugation at 5000 rpm and 4^{o}C for

10 min). The cells were finally resuspended in 1 ml ice cold 0.1 M $CaCl_2$ with 15 % (v/v) glycerol and kept on ice for 10 minutes before they were stored at -80^{o}C in 50 µl aliquots.

3.3 Transformation of competent *E. coli* TB-1- and BL-21-cells

TB-1 cells were transformed for amplification and storage of the pGEV2 vector with PlnJ-GB-1 gene fusion inserted (see section 3.1). For peptide production, BL-21 cells were transformed, with this vector isolated from TB-1 cells (see 3.4),.

50 µl of cells were thawed on ice. Vector DNA, with PlnJ-GB-1 gene fusion insert (see section 3.1) was added, 1/20 $volume_{DNA}/volume_{cells}$. The cells were then incubated for 30 minutes on ice, prior to heat shock, 90 seconds at 42^{o}C. The cells were incubated on ice for 2 minutes before they were transferred to 1 ml SOB media (see Appendix), preheated to 37^{o}C. The cells were incubated, with shaking, for one hour at 37^{o}C.

50 µl of the BL-21 cells were plated on M9 minimal media with agar (see Appendix) in petri dishes. The petri dishes were wrapped in plastic and incubated at 30^{o}C over night. The following day the dishes were inspected for cell growth. Dishes with well defined and isolated cell-colonies were stored at 4^{o}C, wrapped in plastic, for use in for production of peptide.

50 µl of the TB-1 cells were plated LB media with agar (see Appendix) in Petri dishes. After over night incubation at 30^{o}C, 8 well defined isolated cell colonies were transferred to 3 ml LB media each and incubated over night at 37^{o}C with vigorous shaking. After over night incubation, two different 3 ml LB media cultures, containing different clones of the TB-1 cells, were selected randomly amongst the cultures showing cell growth. 1 ml of each of the two were transferred to eppendorf

tubes and 200 µl glycerol was added before the tubes were stored at -80^{0}C to serve as a stock for production of DNA to use in sequencing and transformation of BL-21 for expression.

3.4 Purification of plasmid DNA from *E. coli* TB-1

DNA was isolated using Quiaprep Spin Miniprep (Quiagen) and Nucleo Spin Plasmid (Macherey Nagel) plasmid DNA purification kits.

E. coli TB-1 from the -80^{o}C stock (see 3.3) was grown overnight in 3 ml LB media. A sterile pipette tip was used to scrape a few cells out of the tube of frozen cells and transfer these to the LB media. 1 ml cell culture was centrifuged at 13,000 rpm for 30 seconds in a tabletop centrifuge (see Appendix), the supernatant was discarded. Cells were resuspended in 250 µl Buffer A1 and vortexed. 250 µl Buffer A2 was added and the solution inverted 6-8 times to carefully mix. The solution was incubated at room temperature for one minute before adding 300 µl Buffer A3 and carefully mixing by inverting 6-8 times.

These steps assure lysis of the cells and precipitation of genomic DNA along with remains of the cells. Plasmid DNA is in solution. Careful mixing after adding Buffers A1 and A2 is to avoid shearing of genomic DNA, which would result in genomic DNA in solution and consequent contamination of purified plasmid DNA.

The solution was centrifuged at 13,000 rpm for 10 min and the supernatant was loaded onto a spin column supplied with the kit. The spin column was placed inside a tube designed to collect the flow through and centrifuged at 13,000 rpm. The flow through was discarded and 600 µl Buffer A4 was loaded onto the column. The column was centrifuged twice; first at 13,000 rpm for 1 minute, then at 13,000 rpm for 2 minutes, the flow through was discarded both times.

These steps assure that the plasmid DNA binds to the column, that unbound contaminations are washed of the column and that the column is dry.

50 µl Buffer AE is loaded onto the column and left to incubate for 1 minute at room temperature before placing the column in a clean eppendorf tube and eluting purified plasmid DNA by centrifuging at 13,000 rpm for one minute, collecting the flow through in the eppendorf tube.

In this final step, pure plasmid DNA is eluted from the column.

The above protocol is for the Macherey-Nagel kit, the procedure is identical for the Quiagen kit. The exact composition of the various buffers is not revealed, but the general composition of the buffers is as follows:

- A1 - Resuspension buffer
- A2 – SDS/alkaline lysis buffer
- A3 – Neutralizing buffer, creates appropriate conditions for binding DNA to the silica membrane of the NucleoSpin Plasmid column.
- A4 – Ethanol wash buffer
- AE – elution buffer, 5 mM Tris/HCl, pH (:

3.5 Production and purification of PlnJ

Following transformation, BL-21 cells were grown on M9 agar in Petri dishes. Colonies from the Petri dishes were grown in LB media over night and 5-10 ml were used to inoculate 1 liter of LB media to produce cells for expression of peptide. These large volume cultures were grown to OD 0.6-0.7 prior to harvesting the cells by centrifugation at 6000 rpm (see Appendix) and 4^{o}C for 10 minutes. The cells were washed in 1x M9 salt solution (see Appendix) and resuspended in 250 ml ^{15}N labeled M9 media. Before inducing transcription of the GB-1 PlnJ gene fusion by adding isopropyl-beta-D-thiogalactopyranoside (IPTG) to an end concentration of 1 mM, the cell culture was allowed a one hour growth recovery and clearance of unlabeled

metabolites. Four hours after induction, cells were harvested by centrifugation at 6,000 rpm and 4°C for 10 minutes.

Lysis of the cells was preformed on the Express (see Appendix): The cells were submerged in liquid nitrogen, crushed into a "frozen cell powder" and poured into a metal cylinder. To prevent thawing of the cells, the metal cylinder was kept in a -18°C freezer overnight before use. The cylinder was blocked in the middle by a disc with a small hole in it, about one millimeter, the cylinder was thus divided into two compartments, between which passage is only possible through the small hole in the blocking disc. A piston was entered into the same cylinder-compartment as the cells, and the open end of the other compartment was blocked. The cylinder was placed in a hydraulic press, and 5-6 tons of pressure was applied to the piston, forcing the cells through the small hole in the blocking disc. The cylinder was turned around and the cells pressed back into compartment they were first placed in. The cells were pressed back and forth three times. The cells were lysed mechanically. Ice-crystals in- and out- side the cells sheared the cells as they were pressed through the small hole.

Lysis was followed by resuspension in 20 ml MQ water per one liter LB media culture.

To solubilize proteins in inclusion bodies, guanodine hydrochloride was added to an end concentration of approximately 3 M, and the lysate was incubated in a 50^0C water bath for 1.5 h. To clear insoluble substances from the supernatant, the lysate was centrifuged at 12,500 rpm and 4°C for 20 minutes. To prepare for purification of the target peptide, the supernatant was diluted approximately 5 times in H_2O with 0.1 % tetraflouroacetic acid (TFA), centrifuged at 12,500 rpm (see Appendix) and 4°C for 20 minutes and filtered through a 2 μm syringe filter.

Purification of the fusion protein was performed on the Äkta system (see Appendix). The fusion protein was applied to a ResourceRPC 3 ml column (see Appendix), equilibrated with H_2O and 0.1 % TFA and eluted from the column with a 0 to 100 %

linear 2-propanol containing 0.1 % TFA gradient over 10 column volumes in H_2O with 0.1 % TFA. The flow rate was 1 ml/minute.

Peak fractions containing the fusion protein collected from the ResourceRPC 3ml purification step were treated with cyanogen bromide to cleave the target peptide from its fusion partner prior to purification on a Vydac c18 column. Approximately 47 ml of collected fractions were distributed into 40 eppendorf tubes and dried over night on a vacuum centrifuge (see Appendix). The content of the tubes was resuspended in 1 ml 0.1 M hydrochloric acid each and a small CNBr crystal, about 2 mg, was added.

A rule of thumb states that for CNBr cutting of small peptides, adding equal amounts, by weight, of peptide and CNBr assures the appropriate excess of CNBr. Alternatively, the molar amount of CNBr needed can be determined by multiplying the molar amount of methionines to be cut by 10. The amount of CNBr used in this study exceeded the amounts dictated by both of these methods.

The tubes were wrapped in aluminum foil and left for 12-14 hours before they were opened and the water, CNBr, HCl and volatile side products of the cutting reaction were allowed to evaporate. To speed up the evaporation, a stream of N_2 gas was led over each sample. The content of each tube was resuspended in 850 µl H_2O and left over night to vent off remaining excess CNBr, HCl and volatile side products before it was diluted five times in H_2O with 0.1 % TFA and applied to the Vydac column. The target peptide was eluted from the Vydac reverse phase column with the same buffers and flow rate as on the RescourceRPC columne, but with a gradient to 100% over 20 column volumes.

3.6 Mass spectrometry

PlnJ was expressed as part of a fusion peptide together with the GB1 domain as described above. It was therefore not possible to screen fractions collected during purification for bacteriocin activity until after cleaving of the fusion peptide. A MALDI-TOF type spectrometer (see Appendix) was used to determine the molecular weights of peptides/proteins contained in the different fractions collected during purification. Before cleaving, sinapinic acid was used as matrix, after cleaving α-cyano-4hydroxycinnamic acid was used. The matrix was dissolved in water, 10 mg/ml, and 0.5 µl droplets were applied to a sample plate followed by immediate application of 0.5 µl aliquots of column fractions onto the droplets. The sample/matrix mix was allowed to dry before one spectrum of 50 shots was collected per fraction

3.7 Bacteriocin assay

50 µl of MRS media (see Appendix) were added to all wells on a micro-titer plate. Another 45 µl of MRS media, 2.5 µl of different column fractions and an excess of synthetic plnK (relative to the estimated amount of bacteriocin in the column fractions) was added to all wells in the first column of wells. The approximately 100 µl in the first column of wells were mixed and 50 µl were removed from each well and added to the next column of wells and mixed. Then, 50 µl were removed from that column of wells and added to the next column of wells and mixed. This process was repeated until 50 µl were added to the last column of wells and after mixing, 50 µl were removed from these wells and discarded. The result is two-fold dilution of the sample compared to the previous column of wells. Each row contained a different column fraction combined with synthetic plnK. An indicator strain was added to all wells. This indicator strain consisted of an over-night culture of *Lactobacillus*

plantarum 965 diluted 200 times in MRS media. To serve as a reference for growth inhibition, one well was not added any bacteriocin, only indicator strain and MRS media. The plate was then incubated at 30^{o}C over night before cell densities were determined, relative to the reference well, spectrophotometrically (see Appendix).

3.8 Circular Dichroism spectroscopy

CD measurements (see Appendix) were preformed in a quarts cuvette with a light path of 0.1 cm. The wavelength range was 190-260 nm. Data pitch was 1 nm and response time 2 seconds, while scanning was continuous at 50 nm/min. Protein concentration ranged from 5.0 x 10^{-5} to 4.6 x 10^{-5} molar.

Measurements were performed on peptide in increasing concentrations of DPC to determine an appropriate concentration for NMR studies. To verify that the peptide was unstructured in water, as expected, a DPC free sample was measured. To mimic the conditions in a NMR sample, an aqueous buffer containing 0.1 % TFA was used.

The samples were composed of a volume of aqueous buffer containing 0.1 % TFA to which volumes of peptide and DPC standard solutions were added to reach desired concentrations. The names and compositions of the samples are specified in section 4 Results and discussion, Table 4.2.1, along with results of CD measurements.

The peptide standard solution was made by diluting a sample of the peptide produced for use in NMR studies. The concentration of the peptide sample was determined by UV/visible spectroscopy. Beer-Lamberts Law was used to calculate the concentrations from the spectroscopic measurements. The DPC solution was made by weighing out an amount of dry DPC and dissolving it in an appropriate amount of water.

The data was analyzed with the helical content at the different DPC concentrations in mind. Both the single wavelength method and the full spectrum method were used for

analysis. The single wavelength method uses molar residual elipticity at 222 nm ($[\theta]_{222}$) to calculate the helical content (%α):

$$\%\alpha = 100\% \frac{[\theta]_{222}}{-40000(1-2.5n)}$$

Formula 14

whereas the full spectrum method relies on the CDPro software package [57], which compares the CD spectra with a reference set containing spectra of proteins with known 3D structures.

3.9 NMR

3.9.1 Acquisition and data prosessing

0.47 mg ^{15}N PlnJ, produced and purified as specified above, was dissolved in 270 µl MQ H_2O containing 0.2 mM DSS, 0.1 % TFA and 10 % D_2O. The end concentration of peptide was 0.59 mM. This was the sample used to acquire all 3D NMR spectra (see Appendix).

1.75 mg synthetic PlnJ was dissolved in 600 µl of the same buffer as the ^{15}N peptide sample, giving an end concentration of 1 mM. This sample was used for all the 2D NMR spectra.

To induce structuring, DPC was added to the samples. 10 mg was added to the ^{15}N PlnJ sample, 42 mg to the synthetic PlnJ sample. The difference in the amount of DPC is due to the difference in amount of the produced and the synthetic sample.

The lower volume of the ^{15}N sample required the use of a Shigemi NMR tube (see Appendix), specially designed for low volume samples. The synthetic peptide was studied in a standard 5 mm NMR tube.

The pulse-sequences used to acquire the 3D HSQC-TOCSY and 3D HSQC-NOESY are described in papers by Palmer et al, Kay et al and Sceucher et al [58-60]. Mixing times used were 150 ms for HSQC-NOESY and 80 ms for HSQC-TOCSY. Pulse-sequences used for 2D NOESY and 2D TOCSY are described by Bax et al and Hwang et al [61, 62]. The HSQC pulse-sequence is described by Davis et al [63]. The pulse-sequence for HNHA is described by Vuister and Bax [55], the evolution delay, ζ, was 13.05 ms.

The number of complex data points for the different dimensions of the various spectra are as follows: 2D HSQC; 2048 direct dimension, 32 indirect dimension,

TOCSY and NOESY; 4096 direct dimension, 512 indirect dimension, TOCSY-HSQC; 2048 direct dimension, 21 nitrogen indirect dimension, 160 proton indirect dimension, NOESY-HSQC; 2048 direct dimension, 24 nitrogen indirect dimension, 256 proton indirect dimension.

By adding points with the value zero, the number of complex data points was rounded up to the nearest number N, satisfying $N = 2^n$, where n is any positive integer, and then doubled. After rounding, the total number of complex data points was 2×2^n.

3.9.2 Assignement

The resonances of the nitrogen in the side chains of arginine and tryptophan have been folded in so that they appear in the spectra even though they actually fall outside the sweep width of the spectra.

To assign a protein NMR spectrum is to determine which atom of which amino acid gives rise to which signal.

A 2D nitrogen-proton HSQC spectrum was used to determine the chemical shifts of ^1J coupled nitrogen and proton. Shift values were used to locate N-H-H diagonal peaks in a ^{15}N TOCSY-HSQC spectrum.

Figure 10 below shows areas of the ^{15}N PlnJ HSQC spectrum expected to contain different categories of ^{1}J coupled nitrogen and proton

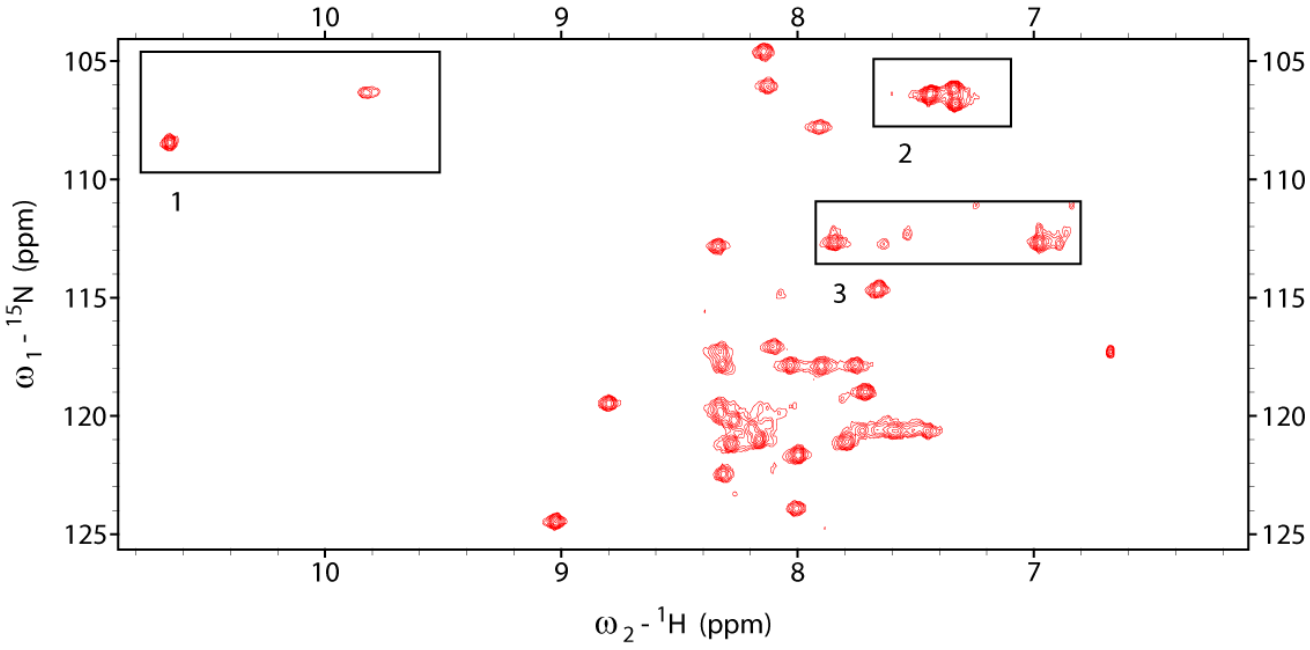

Figure 10 HSQC spectrum showing ^{1}J coupled ^{1}H and ^{15}N of plantaricin J. Squares 1 and 2 show areas containing side chain amines of tryptophan and arginine, respectively. Square 3 contains remaining side chain amines. The rest of the peaks are back bone amid groups. The bulk of amid groups fall in the area below ^{15}N 115 ppm, whereas serines are expected above 115 ppm and glycines above 110 ppm.

Starting from the N-H-H diagonal peak, all peaks belonging to the same spin-system are located. A spin-system is a series of J-coupled spins. Effects of J-couplings across the peptide bond are not detectable, therefore all peaks in a spin-system arises from one amino acid. It is, however, not necessarily so that a spin-system contains all peaks arising from one amino acid; J-couplings between the backbone amid proton of amino acids with aromatic R groups and protons in the aromatic ring are not detectable. In addition, some amino acids give rise to two or more spin-systems, and the spins giving rise to the peaks in each of these may or may not be J-coupled. Arginine contains four nitrogen (see *Figure 5*) and gives rise to four spin-systems in ^{15}N TOCSY-HSQC; these are divided into two pairs, N/N$^{\varepsilon}$ and N$^{\eta 1}$/N$^{\eta 2}$. Within each pair, the proton chemical shifts are the same; i.e. the signals arise from the same

protons, but no proton gives rise to a peak in both of the two pairs of spin-systems since J-couplings are not detectable across the N^{ε}-C^{ζ} bond. It should be noted that some peaks may not be visible in the spectrum, and some proton shifts may be overlapping. Consequently, spin-systems may display less peaks than theoretically possible. On the other hand, overlapping amid proton and nitrogen shifts for two spin-systems will cause these to be displayed on top of each other, appearing to be one spin-system with many peaks.

The amid proton spin-system, all protons with detectable J-couplings to the amid proton, is used as a starting point when deducing which amino acid gives rise to which spin-system. When comparing the theoretical amid spin-systems of all 20 amino acids, 11 spin-system types emerge; *Figure 11* shows these 11 spin-system types.

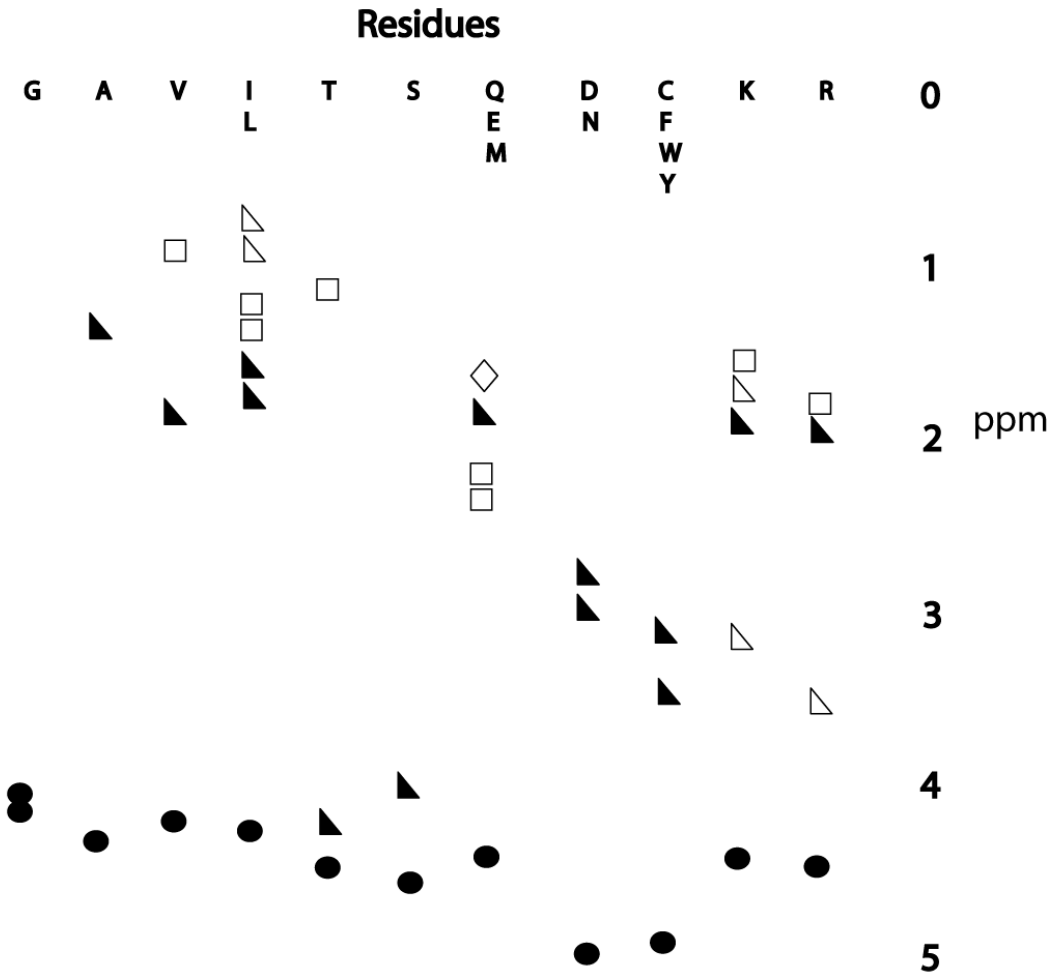

Figure 11 Theoretical fingerprint of the 20 standard amino acids, divided into 11 spin-system types. Filled circles, filled triangles, open squares, open triangles represents H^{α}, H^{β}, H^{γ}, H^{δ}, respectively.

In the lower left centre of the HSQC (see *Figure 10*), there is a peak at approximately ^{15}N-125 ppm, H-9 ppm. Assignment started with this peak. *Figure 12* shows the spin-system associated with this amid group.

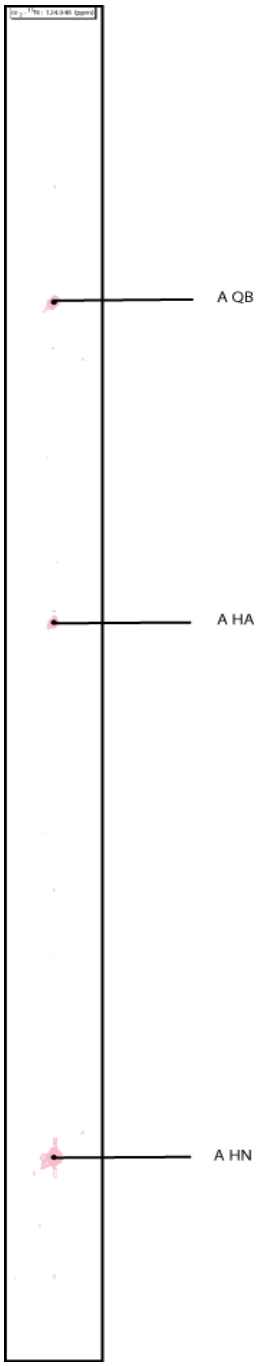

Figure 12 The spin-system of the amid group with chemical shift 9.034-9.034.124.6 (H-H-^{15}N). Comparison with Figure11 shows that this is most likely an alanine.

This spin-system was found to most likely be originating from an alanine when compared to *Figure 11*.

Ones the amid spin-systems have been assigned to a group, sequential assignment starts. A ^{15}N NOESY-HSQC spectrum was overlaid on the ^{15}N TOCSY-HSQC spectrum, that way non-TOCSY type peaks would be easy to see. First, the NOE's between amid protons were mapped. In helical structures, a NOE between H^N of sequential neighbours is expected, see *Table 4*.

Table 4 Expected NOE effects in an α-helical protein

Nuclei sequence	NOE effect
H-H, i-i+1 (H^N-H^N,i,i+1)	Strong
HA-H, i-i+1 (H^α-H^N,i,i+1)	Weak
HA-H, i-i+2 (H^α-H^N,i,i+2)	Medium
HA-H, i-i+3 (H^α-H^N,i,i+3)	Medium
HA-H, i-i+4 (H^α-H^N,i,i+1)	Medium
HB-H, i-i+1 (H^β-H^N,i,i+1)	Medium
H-H, i-i+2 (H^N-H^N,i,i+2)	Medium
HA-HB, i-i+3 (H^α,H^β,i,i+1)	Medium/strong

Figure 13 shows an example of how the NOE's of the above assigned alanine is used to find its sequential neighbor:

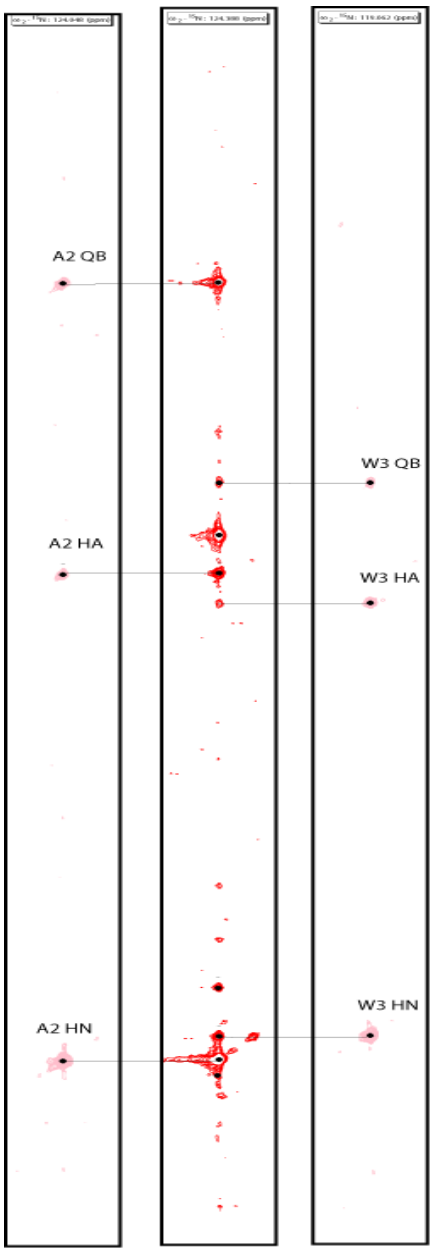

Figure 13 Three spectrum strips showing NOE's between an alanine (later A2) and a tryptophan (later W3). Left and right are the TOCSY-HSQC alanine strip (same as Figure 14) and the TOCSY-HSQC tryptophan strip, respectively, in the middle is the NOESY-HSQC strip of the alanine spin-system. Lines are drawn from the TOCSY-HSQC alanine strip to the NOESY-HSQC alanine strip to show intra residue NOE's, and from the TOCSY-HSQC tryptophan to the NOESY-HSQC alanine to show alanine-tryptophan inter residue NOE's.

When two or three-spin-systems have been linked by inter amid-proton NOE's, the sequence is checked to see if a string of amino acids that gives rise to such linked spin-systems can be found. If such a string only can be found at one place in the sequence, the spin-systems are assigned to these amino acids. The A and the W connected in the above example form a unique sequence string found only in position two and three of the sequence, the sequential assignment A2 and W3 are therefore given.

The easiest way to continue would be to find the sequential neighbors of the already assigned amino acids until the entire peptide was mapped, but even for a small peptide like PlnJ several strings of spin-systems have to be matched against the sequence. Gaps between spin-system sequence matches must be filled by exploring other NOE's, as all the spin-systems form a self consistent web of connectivity's.

The assignment of side chain protons might be a straight forward case of matching the chemical shifts of TOCSY spin-systems against average reported chemical shifts for that amino acid. For other amino acids, like tryptophan, it might involve identifying another spin-system. *Figure 14* shows the assignment of ring-protons of W3 as an example:

In the upper left corner of the HSQC spectrum, *Figure 10*, there is a peak at ^{15}N 108.5 ppm H 10.6 ppm. This peak is positioned in the part of the spectrum where W NE/HE1 ($N^{\varepsilon}/H^{\varepsilon 1}$) HSQC peaks are expected, and consequently checked for NOE connections with W NH (N^{H}) spin-systems in a NOESYHSQC spectrum. A connection with W3 was found, and it was concluded that this was W3 NE/HE1 ($N^{\varepsilon}/H^{\varepsilon 1}$) (see *Figure 14*).

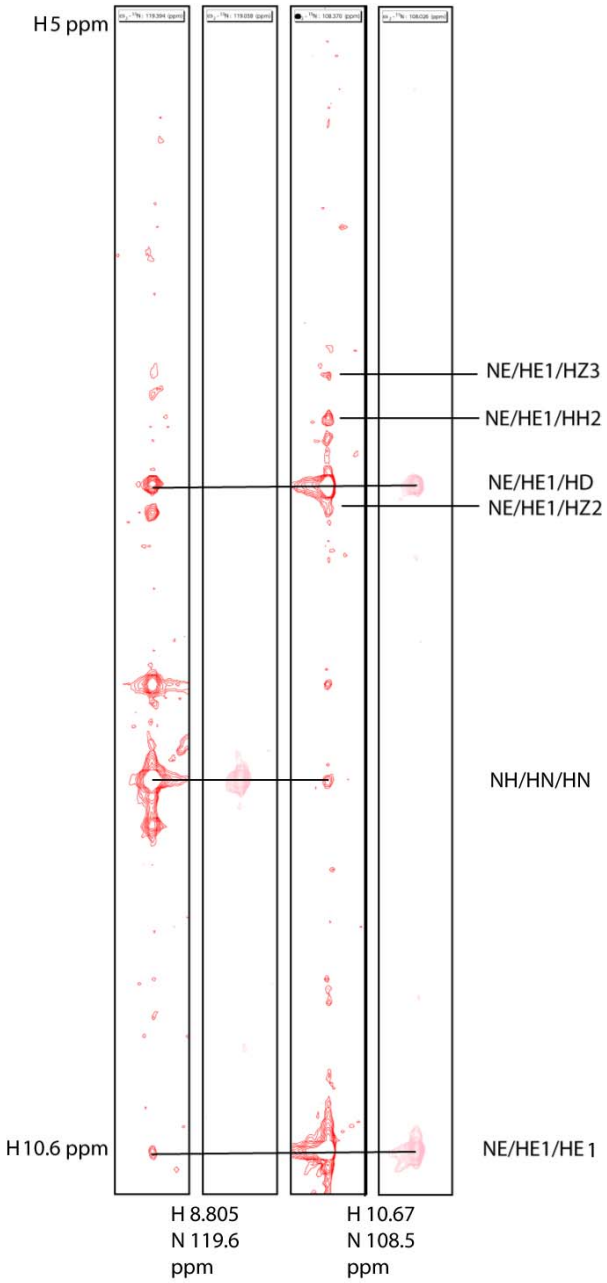

Figure 14 From left to right: W3 NH (N^{H}) spin-system, NOESYHSQC and TOCSYHSQC strips; W3 NE (N^{ε}) spin-system NOESYHSQC and TOCSYHSQC strips. Lines are drawn between the strips to show NOE's connecting the two spin-systems. The HD (H^{δ}) of the ring was identified by being the only proton with a TOCSY connection to HE1 ($H^{\varepsilon 1}$). Lines are drawn out of the strips to mark the remaining ring protons; these were assigned by matching NOE-peaks with peaks belonging to TOCSY-W-ring spin-systems (see Figure 17). W3 HE3 ($H^{\varepsilon 3}$) cannot be seen due to overlap with W3 HD (H^{δ}).

The HE3 ($H^{\varepsilon 3}$), HZ2 ($H^{\zeta 2}$), HZ3 ($H^{\zeta 3}$) and HH2 ($H^{\eta 2}$) protons of the W ring are all ^{3}J-coupled, and consequently form spin-systems in TOCSY. They are, however, not ^{3}J coupled to any proton sitting on a nitrogen and do therefore not show up in the ^{15}N edited TOCSYHSQC. In *Figure 15*, four interconnected 2D TOCSY spin-systems are shown. These where assigned to the W3 HE3 ($H^{\varepsilon 3}$), HZ2 ($H^{\zeta 2}$), HZ3 ($H^{\zeta 3}$) and HH2 ($H^{\eta 2}$) ring protons as they also appear as NOE's in the W3 NE (N^{ε}) NOESYHSQC strip of *Figure 14*.

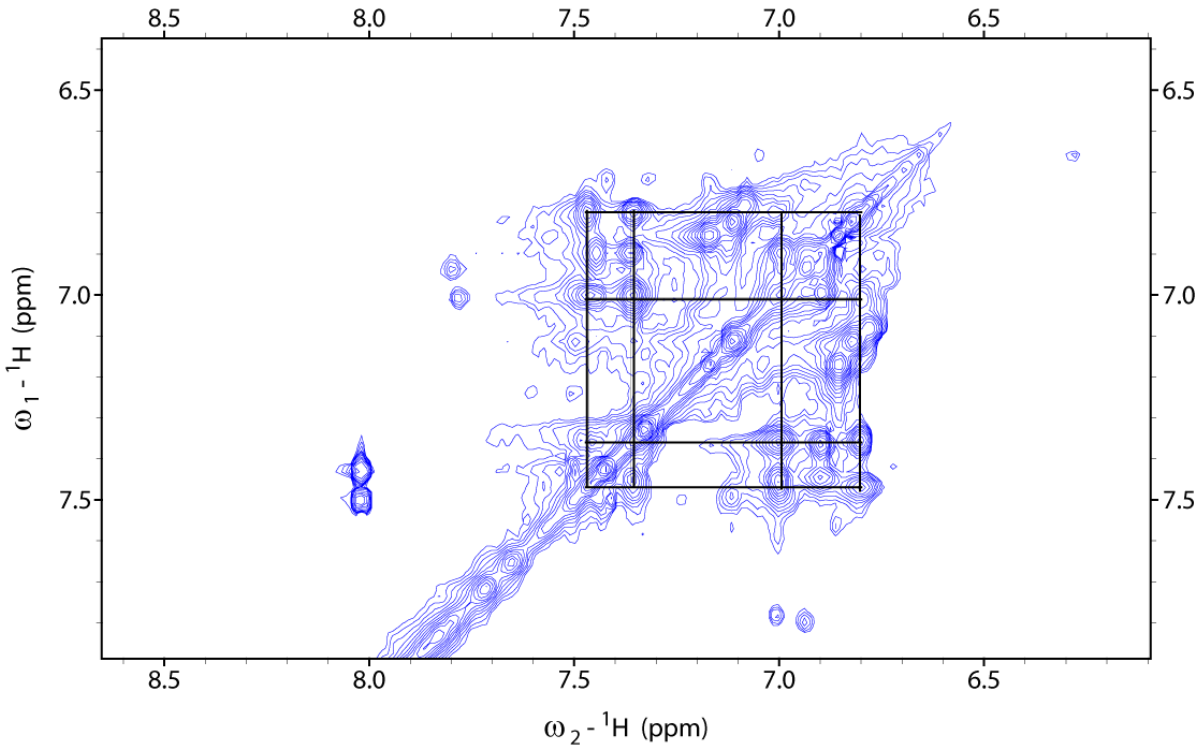

Figure 15 Four interconnected spin-systems in a 2D TOCSY spectrum. Corresponding NOE peaks were found in the NE (N^{ε}) strip of W3, see Figure 16, identifying these as W3 HZ3 ($H^{\zeta 3}$) (6.8 PPM), HH2 ($H^{\eta 2}$) (7 PPM) HZ2 ($H^{\zeta 2}$) (7.4 PPM) and HE3 ($H^{\varepsilon 3}$) (7.5 PPM).

After identifying these as the W3 ring protons, each peak had to be assigned to one of the ring protons. This was achieved by both matching the proton chemical shifts of the peaks against average reported chemical shifts [50], and the relative intensities of the NOE's between the members of the spin-system, see *Figure 16*.

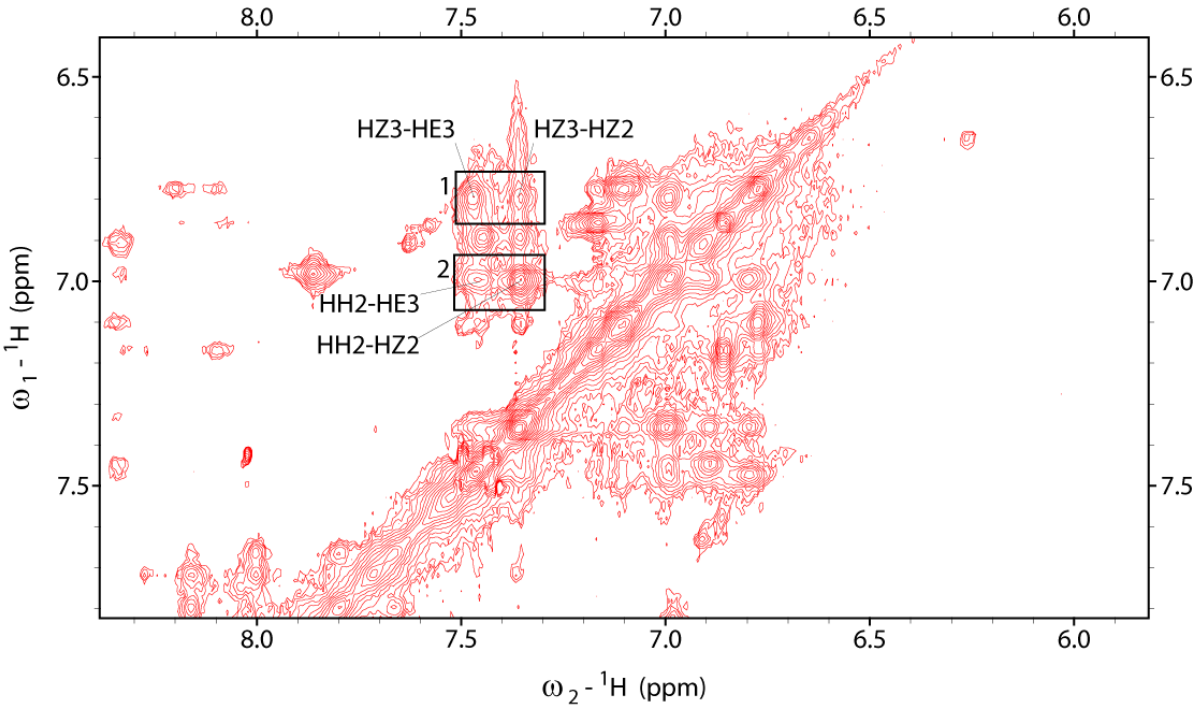

Figure 16 The intraresidual distances of the W3 ring is reflected in its 2D NOESY spectrum: the NOE between HZ3 ($H^{\zeta 3}$) and HE3 ($H^{\varepsilon 3}$) is more intense than the one between HZ3 ($H^{\zeta 3}$) and HZ2 ($H^{\zeta 2}$) due to the smaller distance between HZ3 ($H^{\zeta 3}$) and HE3 ($H^{\varepsilon 3}$), square 1. The distance between HH2 ($H^{\eta 2}$) and HE3 ($H^{\varepsilon 3}$) is greater than between HH2 ($H^{\eta 2}$) and HZ2 ($H^{\zeta 2}$), consequently the HH2 ($H^{\eta 2}$) -HZ2 ($H^{\zeta 2}$) NOE is the most intense, square 2.

3.9.3 Stuctural data

Table 4 lists expected NOE's in an α-helical protein. Most of these involve a proton attached to a nitrogen, these will show up in the NOESYHSQC spectrum used in assignment, and were assigned during this process. The H^{α}-H^{β} i-i+3 will not, therefore the 2D NOESY spectrum used in the assignment of W3 ring protons above, was used to assign these NOE's. Since CYANA will automatically assign NOE's based on input about the chemical shifts of the spins in the amino acids of the protein,

it is crucial to get the assignment of these correct. The NOE cross peaks have to be integrated properly in order for the calibration of them to be correct. Calibration is the process of relating NOE cross peak intensities to the distance between the cross relaxing spins. The intensities of the $H^{N}H^{C\alpha}$ cross peaks were determined through integration, and the $^{3}J_{HNH\alpha}$ was calculated using *Formula 12*.

3.9.4 Cyana

Combined assignment and dYnamics Algorithm for Nmr Applications, Cyana, is a program for automated assignment of NOESY crosspeaks, and structure calculation using a torsion angle dynamics algorithm.

The most important input for CYANA is the Nuclear Overhauser Effects, NOEs, of the molecule. These NOEs have to be assigned, that is, the protons whose dipole moments interact have to be identified. CYANA will do this, provided there is a list of the chemical shifts of the nuclei in the protein. There are, however, always NOEs that cannot be unambiguously assigned. To counter the effects of this, multiple possible assignments are considered for each ambiguous NOE, weighting them according to the inverse of the distance between the protons [64]. Furthermore, all possible assignments are scored according to their network anchoring, which is how well an assignment complies the network of other assignments [64].

Once NOEs are assigned, structure calculations can start.

The structure calculation is based on molecular dynamics, but uses the torsion angles of the protein as the only degrees of freedom. Simulated rises in temperature provide the (virtual) kinetic energy to prevent the structure from getting stuck in a local minimum as the lowest energy structure is sought [64].

Structures provided by the torsion angle dynamics step are evaluated by a target function. Distance constraints based on the assigned NOEs are taken into

consideration. The target function gives a lower score the more torsion angle- and distance- constraints are met [64].

3.9.5 Chemical Shift Indexing

The chemical shift of nuclei in a molecule varies according to the chemical environment they are in. The chemical environment of a nucleus in a given molecule will vary with the structure of the molecule; the orientation of nearby nuclei may vary, and distant nuclei may gain proximity through structural features such as loops. The chemical shift of H^{α} of amino acids varies consistently in accordance with the secondary structure of the protein. When compared to H^{α} of random coil proteins, those of α-helical proteins show an averaged 0.3 ppm lower chemical shift, and those in β-sheet structures show an averaged 0.3 ppm higher shift [65]. Chemical Shift Indexing is assigning amino acids to one of three different groups based on how the chemical shifts of their H^{α} compare to those of the same amino acid in random coil. A chemical shift less than random coil – 0.1 ppm will assign an amino acid to the α-helical group. A chemical shift more than random coil + 0.1 ppm will assign an amino acid to the β-sheet group. In between these two is the random coil group. A string of three or more from one group indicates a possible secondary structure element, or lack of such [66].

3.9.6 TALOS

TALOS is software that uses empirical data to predict phi, φ, and psi, ψ, torsion angles in proteins. These torsion angles are used as angle constraints in further structural calculations. *Figure 17* illustrates the torsion angles predicted by TALOS.

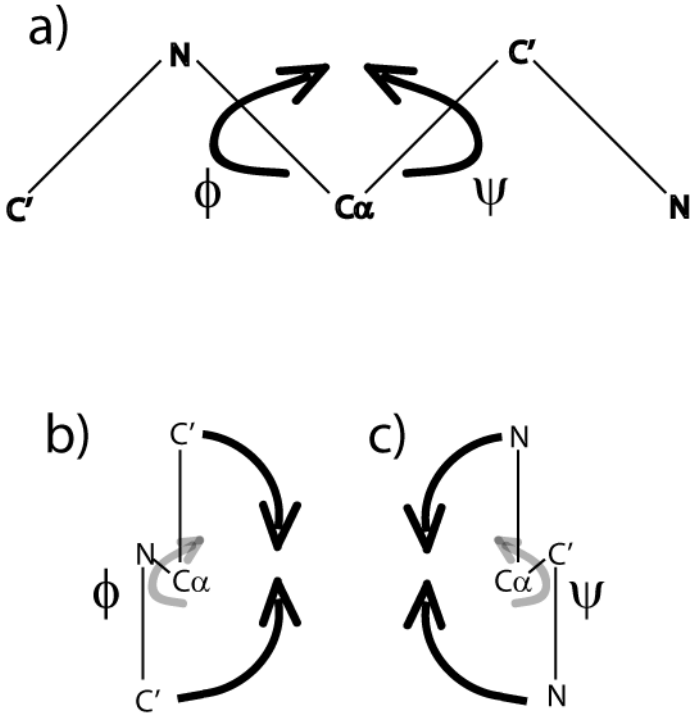

Figure 17 a) The φ and ψ torsion angles represents a rotation about the C^{α}-N and the C^{α}-C bonds, respectively. b) View down the C^{α}-N bond: A change in the φ angle constitutes a change in the distance between the C of two consecutive residues. c) View down the C^{α}-C bond: A change in the ψ angle constitutes a change in the distance between the N of two consecutive residues.

Data from three consecutive amino acids is used to make a prediction about the central of the three. TALOS compares data to that of known structures and predicts based on homology. The data points are H^{α}, C^{α}, C^{β}, C and N chemical shifts, or the ones available. The amino acid type is also taken into consideration; a scoring system biases the prediction towards being based on data from similar sequences. The TALOS data base consists of chemical shifts from 186 proteins [52].

4. Results and discussion

4.1 Production, purification, molecular weight and activity

The yield of PlnJ per liter LB cell culture was approximately 0.1 mg; consequently the procedure had to be scaled up and repeated to produce the amount needed for NMR studies. The final yield of PlnJ was 0.47 mg with molecular weight 2970. Compared to the molecular weight of the non ^{15}N-labled peptide, 2929, this indicates that 99 % of the nitrogen in the produced peptide are ^{15}N. The sample was tested for activity (see section 3.6) and found to be active. *Figures 18* and 19 present chromatograms from the purification process.

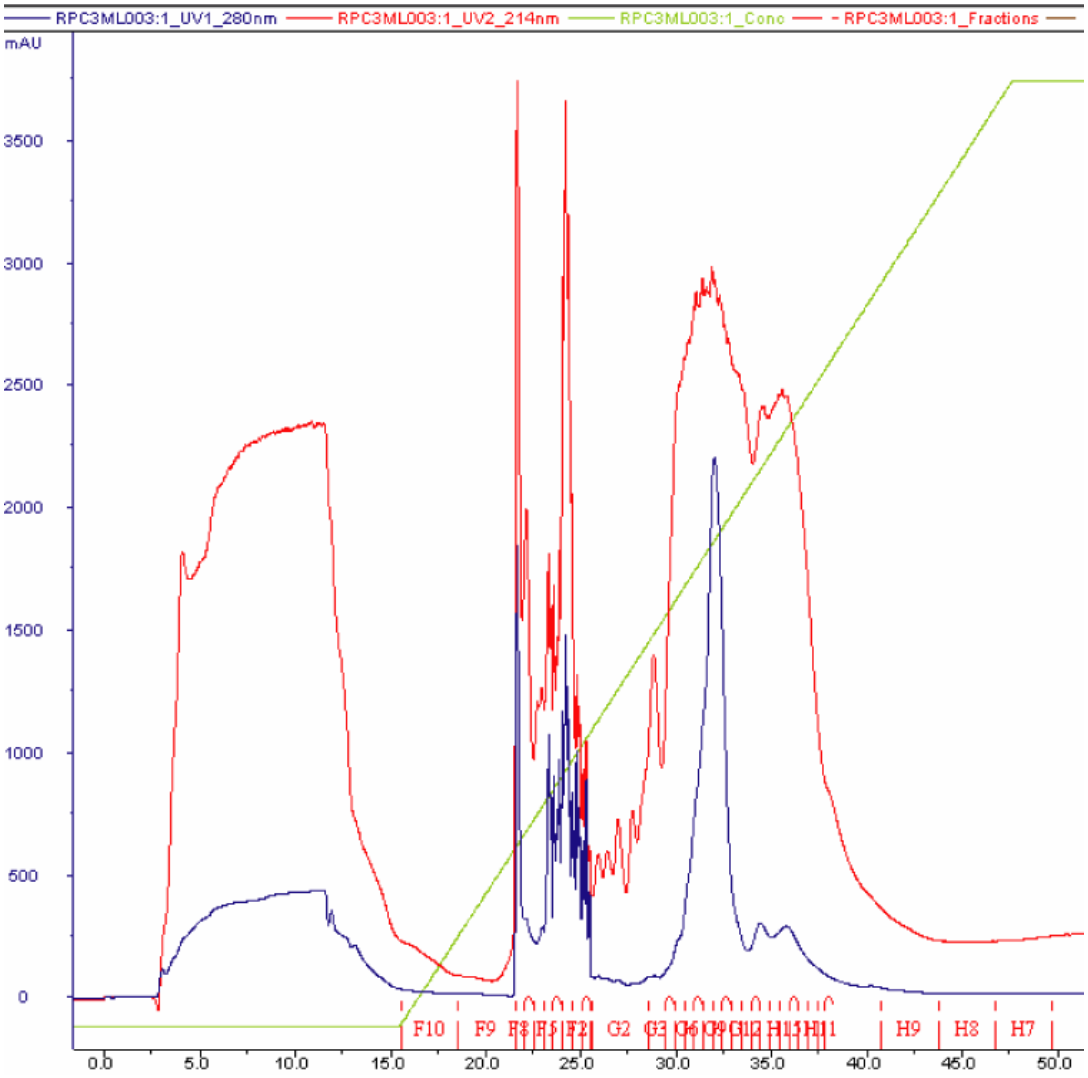

Figure 18 Chromatogram from purification on the RescourceRPC 3ml column. The x-axis is milliliters. The large peak at approximately 32 ml was collected for CNBr cutting and further purification. This chromatogram is representative for one tenth of the peptide production achieved from a one liter inoculated LB media when prepared as described in Section 3.5.

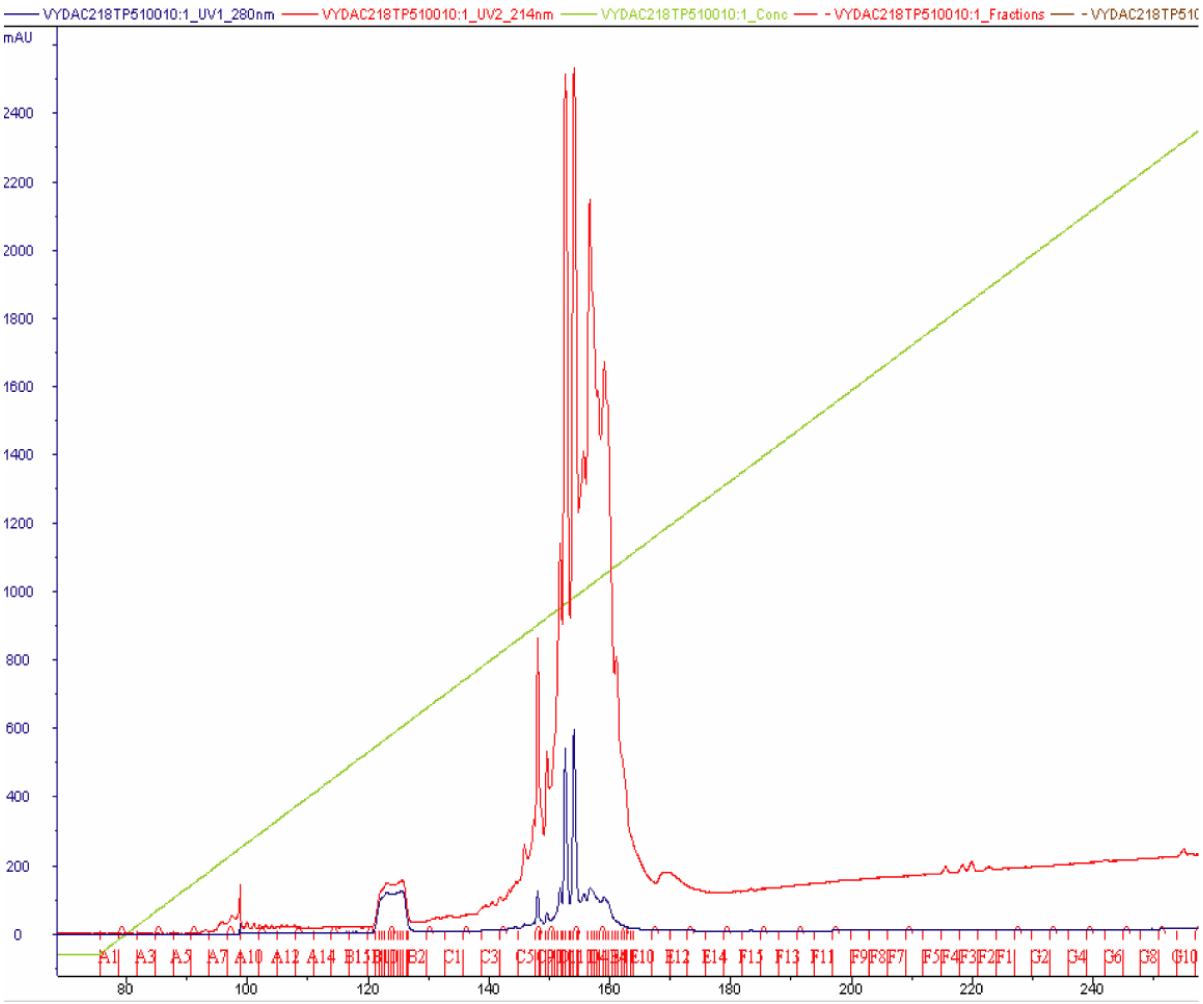

Figure 19 Chromatogram from purification on the Vydac218TP5100 column. X-axis is milliliters. The fastest eluting of the two large peaks, at approximately 152.5 ml, was collected. The slower eluting peak contained a peptide with molecular weight 3010. This chromatogram is representative of the peptide production achieved from four liters of inoculated LB media when prepared as described in section 3.5.

The final 0.47 mg sample was tested for purity on the SMART system (see Appendix) and found to be pure. However, due to computer breakdown, the chromatogram is not available.

4.2 Circular dichroism

To determine the appropriate concentration of DPC in the sample for investigation by NMR, CD spectra of peptide in NMR buffer containing increasing concentrations of DPC were acquired. Results are presented in *Table 5* below. Sample preparation is described in section 2.1 Circular Dichroism. *Figure 20* is a plot of Molar Residue Elipticity throughout the wavelength range of the CD spectra.

Table 5 Sample composition and helical content of PlnJ, determined by use of the single wavelength and full spectrum methods. αd + αr indicates that the helical content calculated with the full spectrum method is the sum of 3_{10}- and α- helical content. The PlnJ0DPC sample was made by combining 130 µl H_2O + 0.1 % TFA and 7 µl 2 mg/ml PlnJ. Subsequent samples were made by adding DPC to this sample.

Sample name	Amount added 0.1M DPC	Final concentrations, M		[DPC] / [PlnJ]	Helical content Single Wavelength method 222 nm	Helical content Full Spectrum method αd + αr
		[DPC]	[PlnJ]			
PlnJ0DPC	0	0	5.0×10^{-5}	-	3.7 %	2 %
PlnJ2DPC	2 µl	1.4×10^{-3}	5.0×10^{-5}	28	8.7 %	5.2 %
PlnJ4DPC	4 µl	2.8×10^{-3}	4.9×10^{-5}	57.1	24.2 %	48.8 %
PlnJ6DPC	6 µl	4.2×10^{-3}	4.8×10^{-5}	87.5	25.6 %	46.5 %
PlnJ8DPC	8 µl	5.5×10^{-3}	4.8×10^{-5}	114.6	26.5 %	45.5 %
PlnJ12DPC	12 µl	8×10^{-3}	4.6×10^{-5}	173.9	25.6 %	43.3 %

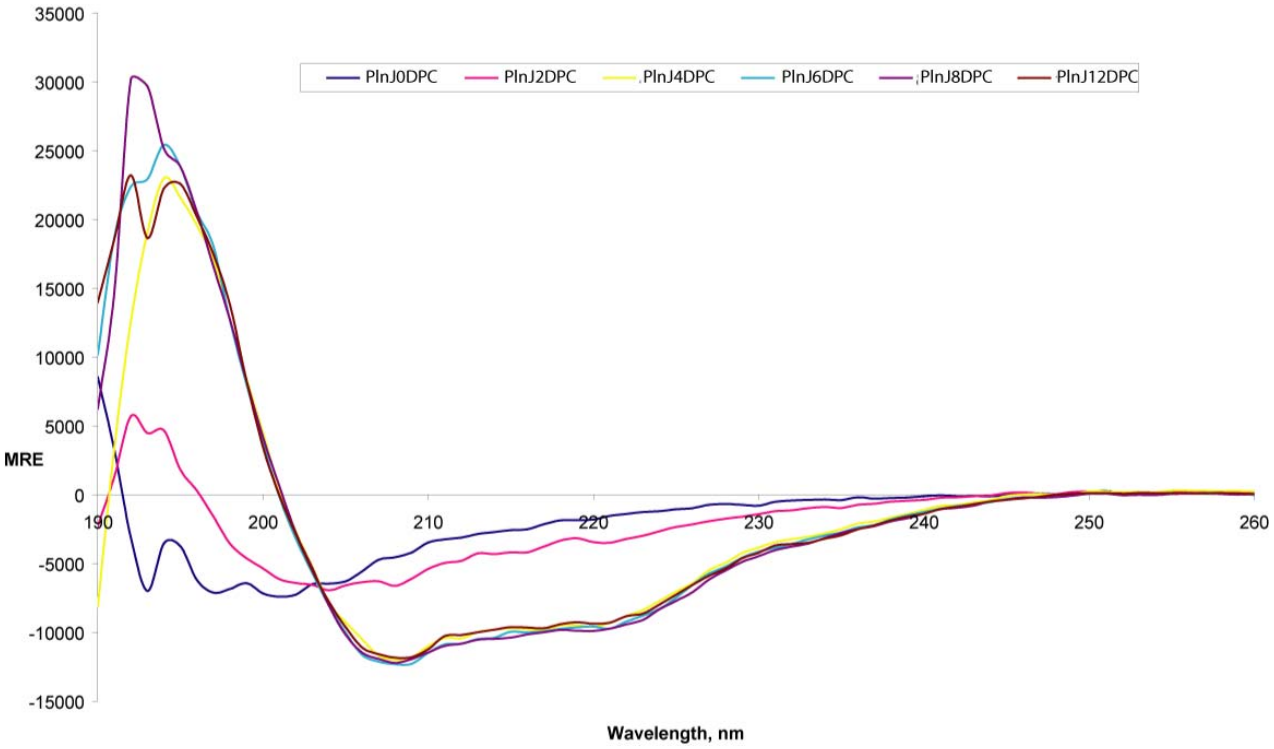

Figure 20 Plot of MRE, molar residue elipticity, as a function of wavelength for samples containing different amounts of DPC. For sample composition see Table 5.

Comparison of *Figure* 3 and *Figure 20* shows that samples 4-12 DPC give CD spectra indicating α-helical structure. The 0 DPC sample spectrum has a random coil shape, whereas the 2 DPC sample spectrum shows changes from random coil which, given the results of the 4-12 DPC samples, might be considered as α-helical structuring. However, on its own the 2 DPC spectrum is ambiguous. Results presented in *Table 4* lead to the same conclusions as *Figure 20*. The findings concur with those of Hauge et al. [18], although the samples of Hauge et al. were composed of peptide and DPC dissolved in 20 mM sodium phosphate buffer instead of H_2O with 0.1 % TFA.

4.3 Chemical Shift Indexing

The amino acids of PlnJ were assigned to three different groups according to the shift of H^{α}. To accommodate graphical presentation, the groups were given the values -1 for α-helical, 0 for random coil, and 1 for β-sheet.

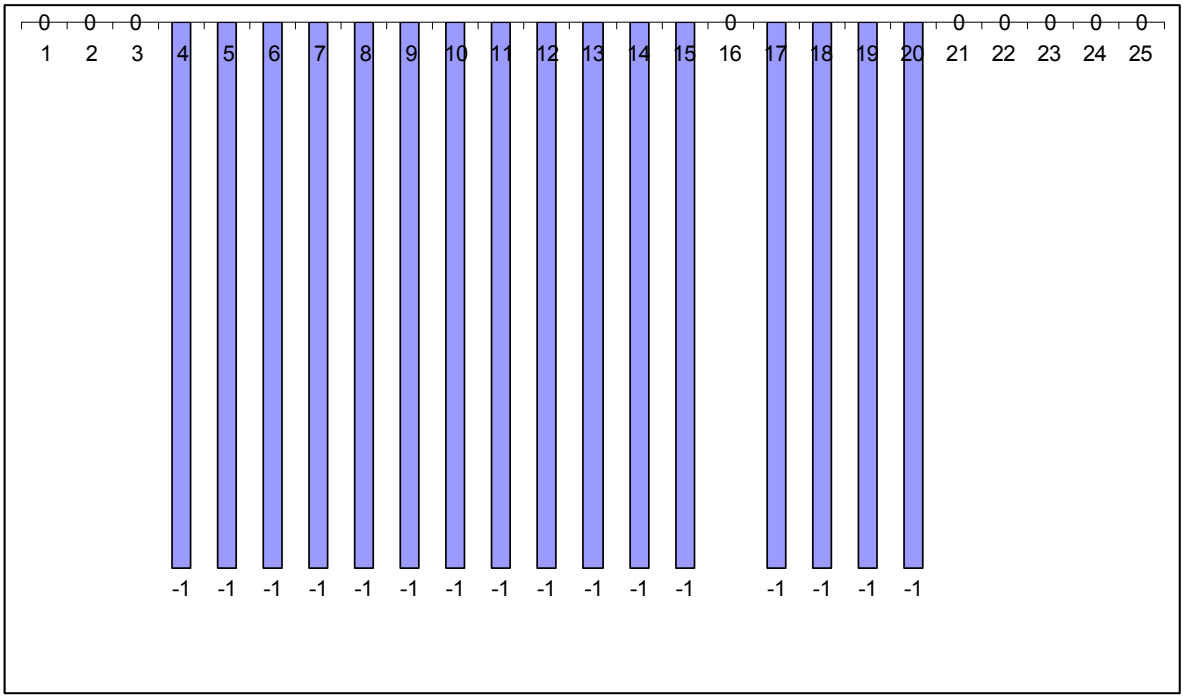

Figure 21 The figure shows that 16, or 64%, of the 25 amino acids of PlnJ are assigned to the α-helical group (-1). In addition to giving a quantitative measure, the results also reveal that the helical content is divided in two groups located at amino acids 4 through 15 and 17 through 20.

Figure 21 shows that 16 out of the 25 amino acids of PlnJ were assigned to the α-helical group. This was as expected and further confirms initial CD results indicating that PlnJ adopts an α-helical structure when exposed to a membrane mimicking substance (DPC).

4.4 TALOS

TALOS dihedral angle predictions are presented as an angle with a standard deviation. These predictions are modified before use as input for the CYANA program; *Table 6* below shows an allowed range for the phi and psi angles of 17 of PlnJ's 25 amino acids, the minimum of the range represents the TALOS output angle minus 2.5 standard deviations, the maximum represents the TALOS output plus 2.5 standard deviations.

Table 6 The table presents modified dihedral angle predictions for 17 of the 25 amino acids of PlnJ made by the TALOS program. Predictions can not be made for the N- and C- terminal amino acids, unambiguous predictions could not be made for the remaining amino acids missing from the list.

Residue Type/number	Angle	Minimum degree	Maximum degree
K 4	PHI	-98.5	-30.1
K 4	PSI	-76.2	-4.9
N 5	PHI	-88.3	-38.3
N 5	PSI	-66.1	-16.1
F 6	PHI	-122.3	-17.1
F 6	PSI	-79.6	-6
W 7	PHI	-86.2	-36.2
W 7	PSI	-70.6	-20.6
S 8	PHI	-88.1	-38.1
S 8	PSI	-67.2	-17.2
S 9	PHI	-88.6	-38.6
S 9	PSI	-70.1	-20.1

L 10	PHI	-87.3	-37.3
L 10	PSI	-68.2	-18.2
R 11	PHI	-85.8	-35.8
R 11	PSI	-68.9	-7.8
K 12	PHI	-132.8	-15.3
K 12	PSI	-77.9	19.8
F 14	PHI	-91.5	-41.5
F 14	PSI	-74.9	2.8
Y 15	PHI	-89.7	-39.7
Y 15	PSI	-64.9	-14.9
D 16	PHI	-120.5	-51.5
D 16	PSI	-68	43.2
E 18	PHI	-92.3	-34.5
E 18	PSI	-87.3	11.3
A 19	PHI	-102.3	-35.2
A 19	PSI	-88.1	20.6
R 21	PHI	-104.9	-33.1
R 21	PSI	-64.2	-7.2
A 22	PHI	-92.5	-42.5
A 22	PSI	-67.4	2.7
I 23	PHI	-123.4	-36.3
I 23	PSI	-66.3	18.9

A perfect α-helix has dihedral angles φ -57and ψ -47, the distorted α-helix, the 3_{10} helix, has angles φ -49 ψ -26 [67]. All of these angles fall within the ranges allowed

by *Table 6*. The TALOS data thus allows the α-helical structure indicated by CD and CSI.

4.5 $^{3}J_{HNH\alpha}$ couplings of PlnJ

$^{3}J_{HNH\alpha}$ coupling constants for 18 of PlnJ's 25 amino acids are presented in *Table 7* below. Coupling constants for glycine residues 1, 13, 17 and 20 were not calculated since glycine has two H^{α} and consequently two coupling constants. Coupling constants are not reported for F6, W7 and L10 since no $H^{N}H^{\alpha}$ cross peaks were found in the spectrum. This might suggest that the coupling constants are so small that the peaks are of lower intensities than the noise of the spectrum, which in turn would make them smaller than the smallest one reported here; R11 1.65.

Table 7 $^{3}J_{HNH\alpha}$ coupling constants, in hertz, of 18 of PlnJ's 25 amino acids.

Residue	$^{3}J_{HNH\alpha}$ coupling constants
G1	-
A2	2.97
W3	3.32
K4	1.95
N5	3.83
F6	-
W7	-
S8	1.85
S9	3.33
L10	-

R11	1.65
K12	3.24
G13	-
F14	4.49
Y15	7.38
D16	3.47
G17	-
E18	2.31
A19	3.29
G20	-
R21	4.93
A22	3.71
I23	6.05
R24	5.9
R25	5.3

A $^{3}J_{HNH\alpha}$ of 3.9 Hz is expected for a perfect α-helix, 4.2 Hz for a 3_{10}-helix, but stretches of three or more amino acids with $^{3}J_{HNH\alpha}$ less than 6 is considered an indication of a helical structure [68]. Apart from Y15 and I 23, all the amino acids have $^{3}J_{HNH\alpha}$ lower than 6 Hz. There is, however, just one stretch of three or more amino acids with $^{3}J_{HNH\alpha}$ lower than 6 Hz; A2 – N5. On the other hand, if one were to assume that amino acids F6, W7 and L10 all had $^{3}J_{HNH\alpha}$ lower than R11 at 1.65 Hz, amino acids A2 through K12 would form a continuous string of $^{3}J_{HNH\alpha}$ lower than 6 Hz.

4.6 Structure calculation of plantaricin J

CYANA returns 20 different structures based on the structural data input. Structures 1 through 20 have a mean target function value (see section 3.9.4) of 0.95, the minimum value is 0.86 (Structure 1), the maximum 1.06 (Structure 20). Non-NOE structural data is presented in sections 4.4 and 4.5. The Upper distance Limit (UPL) list of NOE derived constraints is given in the appendix (section 5.5).

Figure 22 gives three different least square fits of the final 20 structures, *Table 8* gives the root mean square deviation (RMSD) of these fits.

Figure 23 is a ribbon animation of structure 1, showing two helical structure elements separated and flanked by less structured sections composed of turns and bends. The N-terminal helical element runs from residue 3 through 13, the C-terminal helical element runs from residue 18 through 21. In *Figure 24* the non-polar side chains have been colored red, and the polar blue, to display the amphiphilic nature of Structure 1.

Figure 25 is a summary of the NOE derived distance constraints.

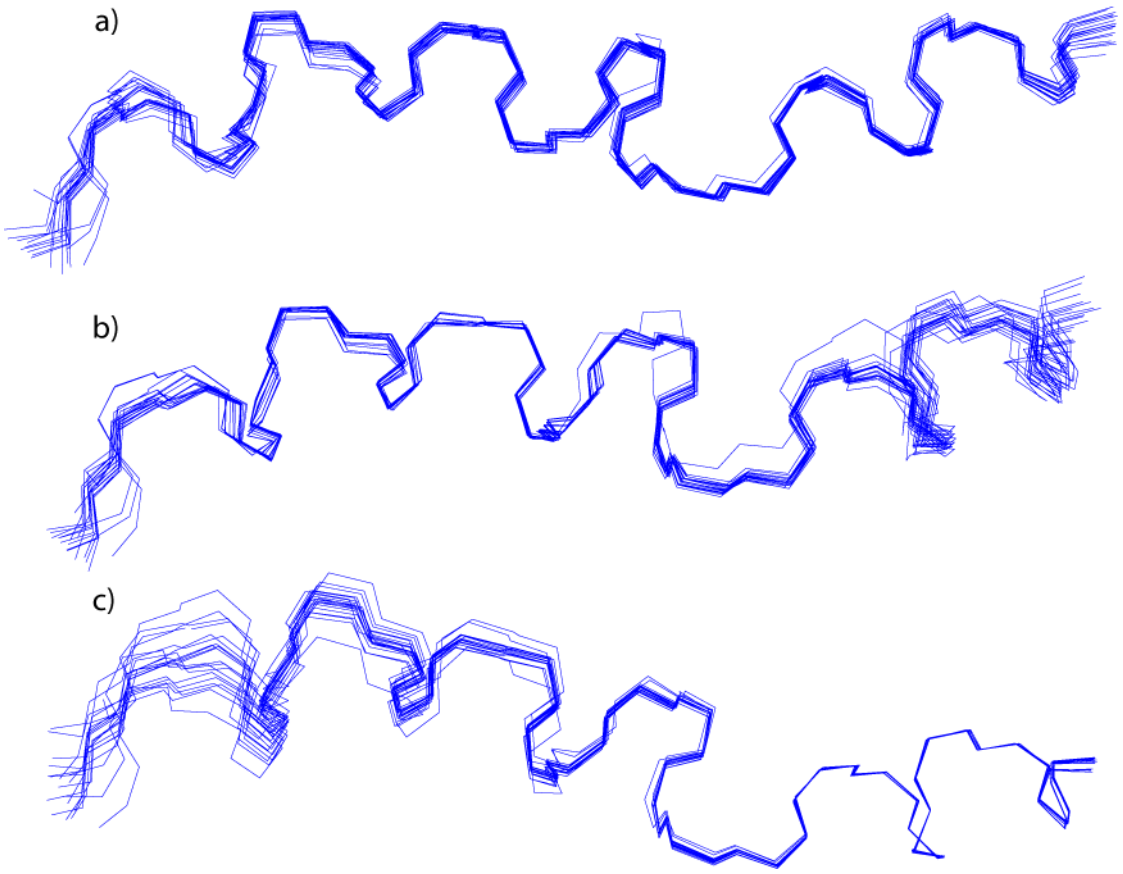

Figure 22 Least square fit of the final 20 structures. a) Least square fit of all 25 residues. b) Least square fit of residues 3-12. c) Least square fit of residues 19-23. Figures made with MOLMOL [69]. The N-terminus is to the left in all three views.

Table 8 Root Mean Square Deviation of fits over different residue ranges. Backbone denotes the continuous chain of carbons and nitrogen connected by the peptide bond, and the carbonyl oxygen.

Fit over residues	RMSD, backbone
1-25	0.47 +/- 0.20 Å
2-12	0.32 +/- 0.20 Å
19-23	0.05 +/- 0.03 Å

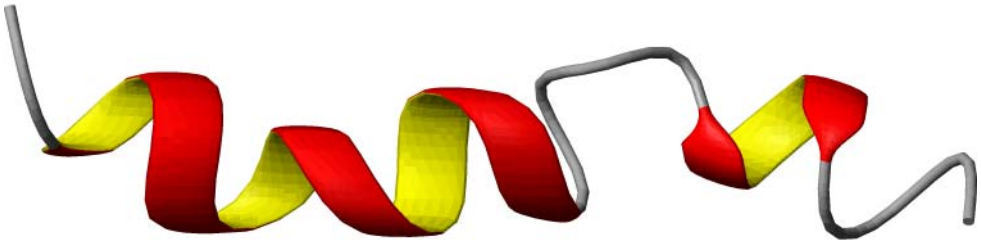

Figure 23 Ribbon animation of Structure 1, N-terminus to the left. The unstructured part between the two helices is residues 14 through 17, the XXXG part of the GXXXG motif. Figure made with MOLMOL [69].

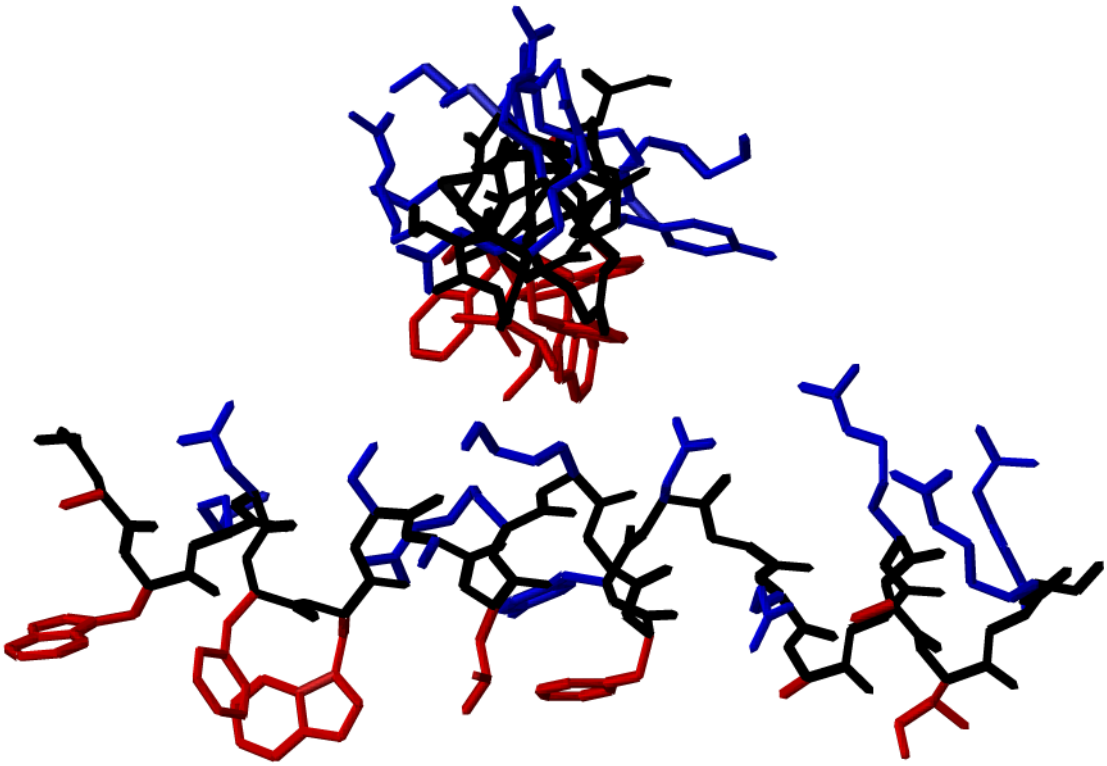

Figure 24 View of the amphiphilic nature of PlnJ. Non-polar side chains of residues A2, W3, F6,W7, L10, F14, A19, A22 and I23 are colored red. Polar side chains of residues K4, N5, S8, S9, R11, K12, Y15, D16, E18, R21, R24 and R25 are colored blue. The view at the top is down the helix, from the N-terminal. The bottom view is from the side, the N-terminus is to the left. Figures made with MOLMOL [69].

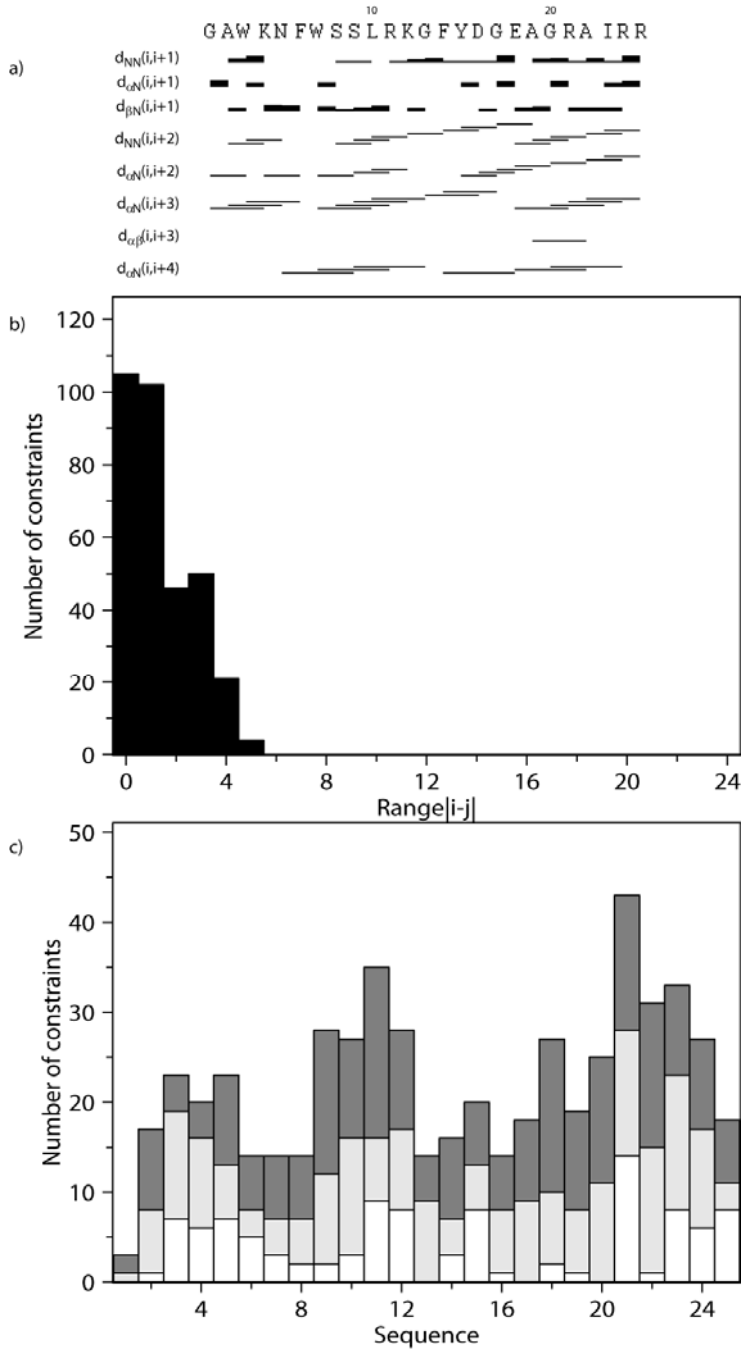

Figure 25 Overview of the NOE derived distance constraints used in the structure calculations of PlnJ. a) gives an overview of which residues are linked by constraints, and also what nuclei in these residues that are linked by constraints. A horizontal line between two sequence positions indicates a constraint. For the sequential constraints, the thickness of the lines is inversely proportional with the square of the upper distance. b) gives an overview of number of constraints per range. c) also gives the number of constraints per range, but as part of a bar giving the number of constraints per residue. The white part of the bar gives the intra-residual constraints, grey sequential constraints and dark grey constraints between residues 2-5 positions apart in the sequence.

4.7 Discussion

Comparing *Figure 25a)* with *Table 4* reveals that most of the expected NOE interactions are present. There are, however, a few points worth mentioning. In α-helixes, the $d_{\alpha\text{-}\beta i,i+3}$ (H^{α}-H^{β},i,i+3) NOEs are expected to be strong, and hence readily observable. Based on *Figure 23*, $d_{\alpha\text{-}\beta i,i+3}$ (H^{α}-H^{β},i,i+3) NOEs are expected in the 3-13 and 18-21 parts of the molecule, eight in total. Only one is reported, between residues A19 and A22. The peaks resulting from these NOEs are expected in the same area of the spectrum as the intra residue- and sequential- H^{α} to H^{β} peaks, which is crowded. Of these eight expected peaks, six are possible, meaning they might be present but overlapping resonances make them difficult to isolate and integrate. The S9H^{α}-K12H^{β} peak was found, but the constraint was deleted by CYANA as it did not affect the structure. The N5H^{α}-S8H^{β} peak was confirmed missing from the spectrum. The sequential $d_{N\text{-}N}$ (H^{N}-H^{N},i,i+1) and $d_{\beta\text{-}N}$ (H^{N}-H^{β},i,i+1) NOEs are lacking between six and ten pairs of residues (respectively). This is in large part due to overlapping resonances, especially the resonances of residues W3, K4, N5, F6 and W7. The sequential H^{α}-H^{N} NOEs are also missing between some residues. Resonance overlap is a reason for this as well, but, because they are generally weaker than the other sequential NOEs, the distance constraints they cause are so generous they have no impact on the structure; consequently CYANA has ignored some of them. Although there are some NOEs missing, all the NOEs present are expected in an α-helical structure. The NOE data alone can thus not be considered proof of an α-helix, but as an indication of one.

Figure 25b) shows that about a third of the NOEs are intra-residue, another third are sequential and the remaining third are medium-range. There are no long range (between residues more than 5 places apart in the sequence) NOEs. This is all in concordance with the suggested α-helical structure.

Figure 25c) is a display of numbers of constraints per residue. The residues with the fewest constraints are all situated in either the N- or C-terminal, or in sequential positions 6-8 or 13-19. These are the unstructured parts of the molecule and the part where connectivity's are difficult to establish due to overlap.

Preliminary investigation, using CD, revealed that PlnJ had an α-helical structure. Chemical Shift Indexing confirmed this, suggesting that there might be a break in the structure around residue 16 and at each end of the molecule. *Figure 23* shows a structure, calculated by CYANA, consisting of two helical regions separated by a GXXXG motif spanning residues G13 to G17 with less structuring at the ends, and thus concur with preliminary findings.

The primary sequence of PlnJ reveals that it can form an amphiphilic helix. *Figure 24* shows that this occurs upon interaction with a membrane mimicking entity such as DPC. With the hydrophobic side facing the membrane, the hydrophilic side would be exposed to the solvent, assuring a thermodynamically favorable positioning of all amino acids. *Figures 22-24* show that PlnJ appears to adopt a structure without pronounced bends that would position parts of the molecule in another plane or orientation than the rest of the molecule. This could indicate that the molecule lies flat along the surface of the micelle. It appears that the N-terminal part, up to W7, could be submerged deeper into the micelle before the polar residues come in conflict with the micelle, as K4 and N5 have their side chains directed away from the hydrophobic side. R11, Y15 and E18 on the other hand have their polar side chains pointed towards the hydrophobic side, limiting how far this part of the molecule could be submerged. This allows PlnJ to be positioned in the micelle such that the extrapolated long axis of the molecule would draw a secant through a thought spherical micelle rather than a tangent. *Figure 26-27* shows both possibilities.

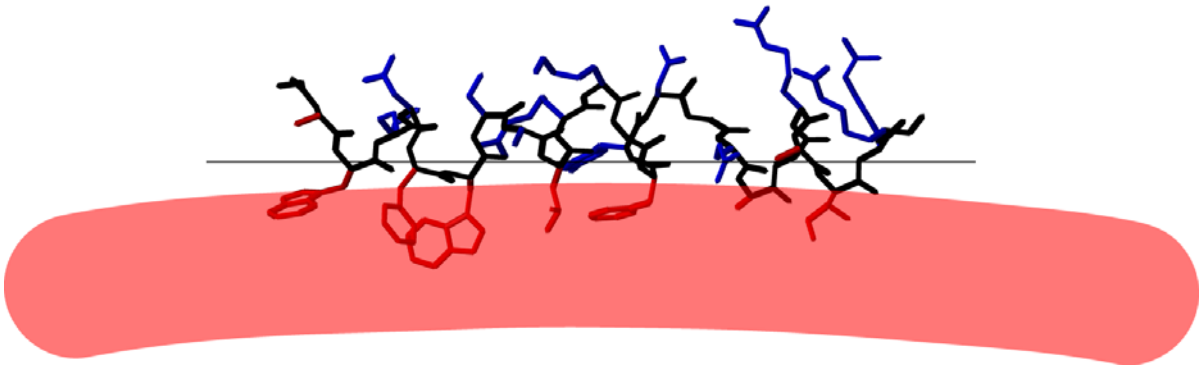

Figure 26 PlnJ embedded in micelle, long axis tangent to the micelle surface. The pink field is the micelle surface. Figure made with MOLMOL [69].

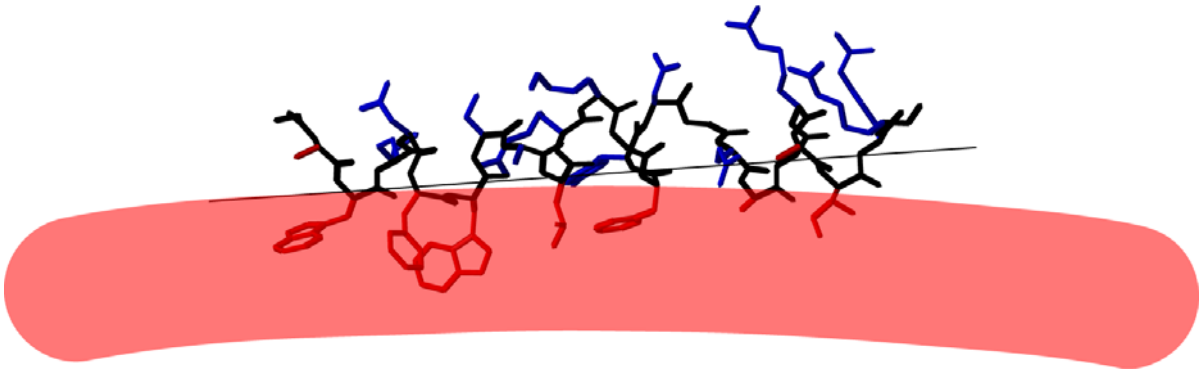

Figure 27 PlnJ embedded in micelle, long axis secant to the micelle surface. The pink field is the micelle surface. Figure made with MOLMOL [69].

In *Figure* 26, W3, F6 and W7 are less deeply embedded in the micelle. The orientation shown in *Figure 27* embeds the aromatic side chains of the N-terminal region deep in the micelle.

It is important to realise that DPC is used as a membrane mimicking structuring agent, not as a model of a target cell. The importance of hydrophobic interactions in peptide folding [67] combined with the fact that PlnJ is unstructured in water suggests that the hydrophobic amino acid residues play a part in the structuring of PlnJ. Their interaction with the DPC micelles resembles that with a cell-membrane and therefore gives a realistic view of the structuring of PlnJ [46]. The two components of plantaricin J/K, PlnJ and PlnK, do, however, not show increased

structuring when combined with DPC together. Thus, DPC micelles can not serve as a system for revealing the interaction of the two peptides, only their individual structures.

Site directed mutagenic studies combined with structural and functional studies have proven useful in revealing the role of individual amino acid residues in the structuring and bactericidal effect of bacteriocins [19, 43, 70]. There has been preformed a mutagenic study on plantaricin J/K showing that the low bulk of the glycine side chains in position 13 and 17 of both peptides is important for the function of the bacteriocin: glycine to alanine mutation are far less detrimental to activity than glycine to lysine mutations (Rogne, unpublished). The other positions in the GXXXG motif are flexible with regards to amino acid side chain bulk and charge, suggesting they do not serve any specific function. It is thus likely that the PlnJ and PlnK peptides of plantaricin J/K interact through the GXXXG motifs, more specifically the glycines, spanning residues 13-17 in both peptides. CD studies of PlnK have revealed that it too adopts an α-helical structure when exposed to DPC micelles [18], which makes it possible that the tertiary structure of plantaricin J/K is similar to that which has been suggested for other two-peptide bacteriocins: two parallel helices [19, 44]. However, if PlnJ and PlnK were to lay in an anti parallel and staggered way, glycine 13 of one peptide interacting with glycine 17 of the other, then the positive charges of the C-terminus of PlnJ and the N-terminus of PlnK would make plantaricin J/K polar, see *Table 9* bottom cell. The mode of action suggested for lactococcin G, relies on a positive charge being pulled through the cell membrane of the target cell, towards the electrically negative inside of the cell [44]. This mode of action would be more plausible in the anti parallel alignment, as a total of 7 positive charges come together at the PlnJ C-term/PlnK N-term part of the molecule, see *Table 9* bottom cell.

Table 9 Distribution of charged amino acid residues in parallel (top) and anti parallel (bottom) interactions between PlnJ and PlnK. Positively charged amino acid residues coloured red, negatively charged blue. The sequences are aligned such that the GXXXG motifs (bold case), which are thought to interact, are opposite each other. It should be noted that this is a sequence alignment, displaying what general part of the molecule (plantaricin J/K) charged amino acid residues are located in, and not a display of relative positioning of amino acids in the structured molecule.

PlnJ N-term	GAWKNFWSSLRK**GFYDG**EAGRAIRR
PlnK N-term	RRSRKNGIGYAI**GYAFG**AVERAVLGGSRDYNK
PlnJ N-term	GAWKNFWSSLRK**GFYDG**EAGRAIRR
PlnK	KNYDRSGGLVAREVA**GFAYG**IAYGIGNKRSRR N-term

Reference:

1. Nissen-Meyer, J. and I.F. Nes, *Ribosomally synthesized antimicrobial peptides: their function, structure, biogenesis, and mechanism of action.* Archives of Microbiology, 1997. **167**(2 - 3): p. 67.
2. Boman, H.G., *Peptide antibiotics and their role in innate immunity.* Annu Rev Immunol, 1995. **13**: p. 61-92.
3. Cammue, B.P., et al., *Gene-encoded antimicrobial peptides from plants.* Ciba Found Symp, 1994. **186**: p. 91-101; discussion 101-6.
4. Sahl, H.G., *Gene-encoded antibiotics made in bacteria.* Ciba Found Symp, 1994. **186**: p. 27-42; discussion 42-53.
5. Jacob, F., et al., *[Definition of some terms relative to lysogeny.].* Ann Inst Pasteur (Paris), 1953. **84**(1): p. 222-4.
6. Tagg, J.R., A.S. Dajani, and L.W. Wannamaker, *Bacteriocins of gram-positive bacteria.* Bacteriol Rev, 1976. **40**(3): p. 722-56.
7. Klaenhammer, T.R., *Genetics of bacteriocins produced by lactic acid bacteria.* FEMS Microbiol Rev, 1993. **12**(1-3): p. 39-85.
8. Ingolf F. Nes, e.a., *Unmodified Peptide-Bacteriocins(CLass II) Produced by Lactic Acid Bacteria*, in *Peptide Antibiotics. Discoverys, Modes of Action and Applications*, M.A.H. Christofer J. Dutton, Hamish A. I. McArthur and Richard G. Wax, Editor. 2001. p. 81-115.
9. Cleveland, J., et al., *Bacteriocins: safe, natural antimicrobials for food preservation.* International Journal Of Food Microbiology, 2001. **71**(1): p. 1-20.
10. Nissen-Meyer, J. and I.F. Nes, *Ribosomally synthesized antimicrobial peptides: their function, structure, biogenesis, and mechanism of action.* Arch Microbiol, 1997. **167**(2/3): p. 67-77.
11. Cotter, P.D., C. Hill, and R.P. Ross, *Bacteriocins: Developing innate immunity for food.* Nature Reviews Microbiology, 2005. **3**(10): p. 777-788.
12. Lozano, J.C.N., et al., *Purification and Amino-Acid-Sequence of a Bacteriocin Produced by Pediococcus-Acidilactici.* Journal of General Microbiology, 1992. **138**: p. 1985-1990.
13. Fimland, G., et al., *Pediocin-like antimicrobial peptides (class IIa bacteriocins) and their immunity proteins: biosynthesis, structure, and mode of action.* Journal of Peptide Science, 2005. **11**(11): p. 688-696.
14. Cotter, P.D., C. Hill, and R.P. Ross, *What's in a name? Class distinction for bacteriocins - Author reply.* Nature Reviews Microbiology, 2006. **4**(2): p. -.
15. Anderssen, E.L., et al., *Antagonistic activity of Lactobacillus plantarum C11: two new two-peptide bacteriocins, plantaricins EF and JK, and the induction factor plantaricin A.* Appl Environ Microbiol, 1998. **64**(6): p. 2269-72.
16. Oppegard, C., et al., *The two-peptide class II bacteriocins: Structure, production, and mode of action.* Journal of Molecular Microbiology and Biotechnology, 2007. **13**(4): p. 210-219.
17. Hauge, H.H., et al., *Amphiphilic alpha-helices are important structural motifs in the alpha and beta peptides that constitute the bacteriocin lactococcin G--enhancement*

of helix formation upon alpha-beta interaction. Eur J Biochem, 1998. **251**(3): p. 565-72.
18. Hauge, H.H., et al., *Membrane-mimicking entities induce structuring of the two-peptide bacteriocins plantaricin E/F and plantaricin J/K.* J Bacteriol, 1999. **181**(3): p. 740-7.
19. Oppegard, C., et al., *Mutational analysis of putative helix-helix interacting GxxxG-motifs and tryptophan residues in the two-peptide bacteriocin lactococcin G.* Biochemistry, 2008. **47**(18): p. 5242-9.
20. Nes, I.F., et al., *Biosynthesis of bacteriocins in lactic acid bacteria.* Antonie Van Leeuwenhoek, 1996. **70**(2-4): p. 113-28.
21. Garneau, S., N.I. Martin, and J.C. Vederas, *Two-peptide bacteriocins produced by lactic acid bacteria.* Biochimie, 2002. **84**(5-6): p. 577-592.
22. Havarstein, L.S., D.B. Diep, and I.F. Nes, *A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export.* Mol Microbiol, 1995. **16**(2): p. 229-40.
23. Holo, H., O. Nilssen, and I.F. Nes, *Lactococcin-a, a New Bacteriocin from Lactococcus-Lactis Subsp Cremoris - Isolation and Characterization of the Protein and Its Gene.* Journal of Bacteriology, 1991. **173**(12): p. 3879-3887.
24. Huhne, K., et al., *Analysis of the sakacin P gene cluster from Lactobacillus sake Lb674 and its expression in sakacin-negative Lb sake strains.* Microbiology-Uk, 1996. **142**: p. 1437-1448.
25. Axelsson, L. and A. Holck, *The Genes Involved in Production of and Immunity to Sakacin-a, a Bacteriocin from Lactobacillus-Sake Lb706.* Journal of Bacteriology, 1995. **177**(8): p. 2125-2137.
26. Allison, G.E. and T.R. Klaenhammer, *Functional analysis of the gene encoding immunity to lactacin F, lafI, and its use as a Lactobacillus-specific, food-grade genetic marker.* Applied and Environmental Microbiology, 1996. **62**(12): p. 4450-4460.
27. Quadri, L.E.N., et al., *Characterization of the Protein Conferring Immunity to the Antimicrobial Peptide Carnobacteriocin B2 and Expression of Carnobacteriocins B2 and Bm1.* Journal of Bacteriology, 1995. **177**(5): p. 1144-1151.
28. Venema, K., et al., *Functional-Analysis of the Pediocin Operon of Pediococcus-Acidilactici Pac1.0/Pedb Is the Immunity Protein and Pedd Is the Precursor Processing Enzyme.* Molecular Microbiology, 1995. **17**(3): p. 515-522.
29. Dayem, M.A., et al., *The putative immunity protein of the Gram-positive bacteria Leuconostoc mesenteroides is preferentially located in the cytoplasm compartment.* Fems Microbiology Letters, 1996. **138**(2-3): p. 251-259.
30. Flynn, S., et al., *Characterization of the genetic locus responsible for the production of ABP-118, a novel bacteriocin produced by the probiotic bacterium Lactobacillus salivarius subsp salivarius UCC118.* Microbiology-Sgm, 2002. **148**: p. 973-984.
31. Diep, D.B., L.S. Havarstein, and I.F. Nes, *Characterization of the locus responsible for the bacteriocin production in Lactobacillus plantarum C11.* J Bacteriol, 1996. **178**(15): p. 4472-83.
32. Diep, D.B., L.S. Havarstein, and I.F. Nes, *A bacteriocin-like peptide induces bacteriocin synthesis in Lactobacillus plantarum C11.* Molecular Microbiology, 1995. **18**(4): p. 631-639.

33. Hauge, H.H., et al., *Plantaricin A is an amphiphilic alpha-helical bacteriocin-like pheromone which exerts antimicrobial and pheromone activities through different mechanisms.* Biochemistry, 1998. **37**(46): p. 16026-32.
34. Diep, D.B., et al., *Inducible bacteriocin production in Lactobacillus is regulated by differential expression of the pln operons and by two antagonizing response regulators, the activity of which is enhanced upon phosphorylation.* Molecular Microbiology, 2003. **47**(2): p. 483-494.
35. Risoen, P.A., et al., *Identification of the DNA-binding sites for two response regulators involved in control of bacteriocin synthesis in Lactobacillus plantarum C11.* Molecular and General Genetics, 1998. **259**(2): p. 224-232.
36. Risoen, P.A., et al., *Regulation of bacteriocin production in Lactobacillus plantarum depends on a conserved promoter arrangement with consensus binding sequence.* Molecular Genetics and Genomics, 2001. **265**(1): p. 198-206.
37. Johnsborg, O., D.B. Diep, and I.F. Nes, *Structural analysis of the peptide pheromone receptor PlnB, a histidine protein kinase from Lactobacillus plantarum.* Journal of Bacteriology, 2003. **185**(23): p. 6913-6920.
38. Maldonado, A., R. Jimenez-Diaz, and J.L. Ruiz-Barba, *Induction of plantaricin production in Lactobacillus plantarum NC8 after coculture with specific gram-positive bacteria is mediated by an autoinduction mechanism.* Journal of Bacteriology, 2004. **186**(5): p. 1556-1564.
39. Kleerebezem, M. and L.E. Quadri, *Peptide pheromone-dependent regulation of antimicrobial peptide production in Gram-positive bacteria: a case of multicellular behavior.* Peptides, 2001. **22**(10): p. 1579-1596.
40. Kleerebezem, M., et al., *Quorum sensing by peptide pheromones and two-component signal-transduction systems in Gram-positive bacteria.* Molecular Microbiology, 1997. **24**(5): p. 895-904.
41. Garneau, S., et al., *Purification and characterization of brochocin A and brochocin B(10-43), a functional fragment generated by heterologous expression in Carnobacterium piscicola.* Appl Environ Microbiol, 2003. **69**(3): p. 1352-8.
42. Senes, A., D.E. Engel, and W.F. DeGrado, *Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs.* Curr Opin Struct Biol, 2004. **14**(4): p. 465-79.
43. Oppegard, C., et al., *Analysis of the two-peptide bacteriocins lactococcin G and enterocin 1071 by site-directed mutagenesis.* Appl Environ Microbiol, 2007. **73**(9): p. 2931-8.
44. Rogne, P., et al., *Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin lactococcin G.* Biochim Biophys Acta, 2007.
45. Fimland, N., *Strukturbestemmelse av to-peptidbakteriosinet plantaricin EF i membranlignende miljø ved hjelp av NMR-spektroskopi*, in *Kjemisk Institutt*. 2007, Universitetet i Oslo: Oslo.
46. Beswick, V., et al., *Dodecylphosphocholine micelles as a membrane-like environment: new results from NMR relaxation and paramagnetic relaxation enhancement analysis.* Eur Biophys J, 1999. **28**(1): p. 48-58.
47. Laidler, K.J.M., John H; Sanctuary, Bryan C, *Physical chemistry*. 4th ed. 2003: Houghton Mifflin. 1060.
48. Cavanagh, J.F., W.J.; Palmer, A.G.; Rance, M.; Skelton, N.J., *Protein NMR Spectroscopy*. 2nd ed. 2007: Elsevier Academic Press. 885.

49. Claridge, T.D.W., *High-Resolution NMR Techniques in Organic Chemistry* First ed. Tetrahedron Organic Chemistry Series. Vol. 19. 1999: Elsevier. 382.
50. Ulrich, E.L., et al., *BioMagResBank.* Nucleic Acids Res, 2008. **36**(Database issue): p. D402-8.
51. Wishart, D.S., B.D. Sykes, and F.M. Richards, *Relationship between Nuclear-Magnetic-Resonance Chemical-Shift and Protein Secondary Structure.* Journal of Molecular Biology, 1991. **222**(2): p. 311-333.
52. Cornilescu, G., F. Delaglio, and A. Bax, *Protein backbone angle restraints from searching a database for chemical shift and sequence homology.* J Biomol NMR, 1999. **13**(3): p. 289-302.
53. Markley, J.L., et al., *Recommendations for the presentation of NMR structures of proteins and nucleic acids--IUPAC-IUBMB-IUPAB Inter-Union Task Group on the standardization of data bases of protein and nucleic acid structures determined by NMR spectroscopy.* Eur J Biochem, 1998. **256**(1): p. 1-15.
54. Wishart, D.S., et al., *1H, 13C and 15N chemical shift referencing in biomolecular NMR.* J Biomol NMR, 1995. **6**(2): p. 135-40.
55. Vuister, G.W. and A. Bax, *Quantitative J Correlation - a New Approach for Measuring Homonuclear 3-Bond J(H(N)H(Alpha) Coupling-Constants in N-15-Enriched Proteins.* Journal of the American Chemical Society, 1993. **115**(17): p. 7772-7777.
56. Huth, J.R., et al., *Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR.* Protein Science, 1997. **6**(11): p. 2359-2364.
57. Sreerama, N. and R.W. Woody, *Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set.* Analytical Biochemistry, 2000. **287**(2): p. 252-260.
58. Kay, L.E., P. Keifer, and T. Saarinen, *Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity.* Journal of the American Chemical Society, 1992. **114**(26): p. 10663-10665.
59. Palmer, A.G., et al., *Sensitivity Improvement in Proton-Detected 2-Dimensional Heteronuclear Correlation Nmr-Spectroscopy.* Journal of Magnetic Resonance, 1991. **93**(1): p. 151-170.
60. Schleucher, J., et al., *A General Enhancement Scheme in Heteronuclear Multidimensional Nmr Employing Pulsed-Field Gradients.* Journal of Biomolecular Nmr, 1994. **4**(2): p. 301-306.
61. Bax, A. and D.G. Davis, *Mlev-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy.* Journal of Magnetic Resonance, 1985. **65**(2): p. 355-360.
62. Hwang, T.L. and A.J. Shaka, *Water Suppression That Works - Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients.* Journal of Magnetic Resonance Series A, 1995. **112**(2): p. 275-279.
63. Davis, A.L., et al., *Experiments for Recording Pure-Absorption Heteronuclear Correlation Spectra Using Pulsed Field Gradients.* Journal of Magnetic Resonance, 1992. **98**(1): p. 207-216.
64. Guntert, P., *Automated NMR structure calculation with CYANA.* Methods Mol Biol, 2004. **278**: p. 353-78.

65. Williamson, M.P., *Secondary-Structure Dependent Chemical-Shifts in Proteins.* Biopolymers, 1990. **29**(10-11): p. 1428-1431.
66. Wishart, D.S., B.D. Sykes, and F.M. Richards, *The Chemical-Shift Index - a Fast and Simple Method for the Assignment of Protein Secondary Structure through Nmr-Spectroscopy.* Biochemistry, 1992. **31**(6): p. 1647-1651.
67. Voet, D., Voet, J.G., *Biochemistry*. 2nd ed. 1995: John Wiley & Sons, Inc.
68. Wuthrich, K., *NMR of proteins and nucleic acids*. 1986: Jon Wiley & Sons, Inc.
69. Koradi, R., M. Billeter, and K. Wuthrich, *MOLMOL: a program for display and analysis of macromolecular structures.* J Mol Graph, 1996. **14**(1): p. 51-5, 29-32.
70. Fimland, G., V.G. Eijsink, and J. Nissen-Meyer, *Mutational analysis of the role of tryptophan residues in an antimicrobial peptide.* Biochemistry, 2002. **41**(30): p. 9508-15.

5. Appendix

5.1 Hardware

Optical density measurements: Shimadzu UV 160 A spectrophotometer.

Centrifugation 6,000 rpm: Beckman Coulter J2-MC, JA 10 rotor.

Centrifugation 12,500 rpm: Beckman Coulter Avanti J-20 XP, JA 25.50 rotor.

Tabletop centrifuge 13,000 rpm: Kendro Heraeus Biofuge Pico, 24x1.5 ml place rotor.

Cell lysis: Express, AB Biox.

Chromatography: ÄKTA *purifier*; pump P-900, UV-metre UV-900, fraction collector Frac-950, injection valve INV-907, mixer M-925, Unicorn software. Amersham biosciences

Chromatographic columns: RescourceRPC 3 ml, 15µm polystyrene/divinylbenzene ads, Amersham biosciences. Vydac 218TP510, n-octadecyl reverse phase based on 300 Å silica, Grace Vydac.

Mass spectrometry: Voyager-DE RP MALDI-TOF spectrometer, PerSeptive Biosystems.

Vaccume centrifuge: Saveant Speed Vac concentrator, Savant Instruments Inc.

Bacteriocin assay: microplate reader, SUNRISE

Circular dichroism: Jasco J-810 spectropolarimeter

NMR spectroscopy: 600 MHz Bruker Avance II spectrometer. Shigemi tubes, Sigma Aldrich

5.2 Chemicals

Acetonitril	Merck
α-cyano-4-hydroxycinnamic acid	Sigma Aldrich
Agar	Merck
Ampicilin	Calbiochem
Biotin	Merck
$CaCl_2$	Merck
Chloramphenicol	Sigma Aldrich
CNBr	Sigma Aldrich
$CoCl_2$ * 6 H_2O	Merck
Guanodine hydrochloride	Sigma Aldrich
H_3BO_3	Merck
IPTG	VWR
KCl	Merck
$MgCl_2$	AppliChem
$MgSO_4$	Merck

$MnCl_2 * 4 H_2O$	Merck
MRS	Oxoid
NaCl	AnalR
NaH_2PO_4	Merck
Na_2HPO_4	Merck
$^{15}NH_4Cl$	Sigma Aldrich, Isotec
2-propanol	Merck
3, 5-Dimethoxy-4-hydroxycinnamic acid (Sinapinic acid)	Sigma Aldrich
Sodium 2, 2-dimethyl-2-silapentan-5-sulphonate (DSS)	Larodan AB
Thiamin	Sigma Aldrich
Trifluoracetic acid	Sigma Aldrich
Tryptone	Merck
Yeast extract	Merck
$ZnSO_4 * 7 H_2O$	Merck

5.3 Media and solutions

M9 minimal media

Per litre:

100 ml	M9 salts (10x)
10 ml	trace element solution (100x)
20 ml	20% (w/v) glucose
0.5 g	NH_4Cl
1 ml	1M $MgSO_4$
300 µl	1M $CaCl_2$
2 ml	0.5 mg/ml biotin
1 ml	1mg/ml thiamin
	Antibiotics

Ingredients, exept glucose, biotin, thiamin and antibiotics were dissolved in 977 ml MQ H_2O and autoclaved for 20 minutes at 121^{o}C. Then the remaining ingredients were added. Glucose had been autoclaved on its own, biotin, thiamin and antibiotic had been passed through a 0.5 µm syringe-filter. Antibiotics used were ampicilin, end concentration 100 µg/ml, and chloramphenicol, end concentration 50 µg/ml. To make plates, 1.5 % agar was added before the media was poured onto Petri-dishes.

M9 salts (10x)

Per litre:

60 g	Na_2HPO_4
30 g	KH_2PO_4
5 g	NaCl

Dissolve in 1000 ml MQ H_2O, autoclave for 20 minutes at 121^{o}C.

Trace element solution (100x)

Per litre:

5 g	EDTA
830 mg	$FeCl_3$ * 6 H_2O
177 mg	$ZnSO_4$ * 7 H_2O
19 mg	$CuSo_4$ * 2 H_2O
10 mg	$CoCl_2$ * 6 H_2O
10 mg	H_3BO_3
1.3 mg	$MnCl_2$ * 4 H_2O

Dissolve EDTA in 750 ml MQ H_2O, adjust pH to 7.5 with NaOH. Dissolve rest of ingredients, add MQ H_2O until total volume is 1000 ml.

SOB media

Per litre:

20 g	tryptone
5 g	yeast extract
2 ml	5M NaCl
2.5 ml	1M KCl
10 ml	1M $MgCl_2$
10 ml	1M $MgSO_4$

Dissolv in < 1000ml MQ H_2O, adjust to 1000 ml. Autoclave 20 minutes at 121^{o}C.

LB media

Per litre:

10 g	NaCl
10 g	tryptone
5 g	yeast exstract

Dissolve in < 1000ml MQ H_2O, adjust to 1000 ml. Autoclave 20 minutes at 121^{o}C. Antibiotics used were ampicilin 100 µg/ml (*E. coli* TB-1 and BL-21) and chloramphenicol 50 µg/ml (*E. coli* BL-21). To make plates, 1.5 % agar was added before the media was poured onto Petri-dishes.

MRS

Ready made, mix only (see Appendix, Chemicals). Autoclave at 121^{o}C for 20 minutes.

5.4 Chemical shifts of PlnJ

CYANA nomenclature.

Group	Atom	Nuc	Shift
G1	QA	1H	3.889
A2	HA	1H	4.261
A2	HN	1H	9.035
A2	N	15N	124.565
A2	QB	1H	1.409
W3	HA	1H	4.558
W3	HD1	1H	7.327
W3	HE1	1H	10.671
W3	HE3	1H	7.462
W3	HN	1H	8.805
W3	N	15N	119.574
W3	QB	1H	3.380
K4	HA	1H	4.109
K4	HG2	1H	1.759
K4	HG3	1H	1.550
K4	HN	1H	8.334
K4	N	15N	120.002

K4	QB	1H	1.934
K4	QD	1H	1.461
N5	HA	1H	4.563
N5	HD21	1H	7.853
N5	HD22	1H	6.986
N5	HN	1H	8.326
N5	N	15N	117.887
N5	ND2	15N	112.756
N5	QB	1H	2.919
F6	HA	1H	4.290
F6	HB2	1H	3.232
F6	HB3	1H	3.154
F6	HD1	1H	6.984
F6	HD2	1H	7.075
F6	HE1	1H	7.469
F6	HE2	1H	7.855
F6	HN	1H	8.325
F6	HZ	1H	6.874
F6	N	15N	122.552
W7	HA	1H	4.027
W7	HB2	1H	3.214

W7	HB3	1H	3.367
W7	HE3	1H	7.442
W7	HN	1H	8.356
W7	HZ2	1H	7.092
W7	HZ3	1H	6.876
W7	N	15N	112.936
S8	HA	1H	4.284
S8	HN	1H	8.340
S8	N	15N	119.781
S8	QB	1H	4.108
S9	HA	1H	4.156
S9	HB2	1H	3.969
S9	HB3	1H	3.747
S9	HN	1H	7.906
S9	N	15N	117.975
L10	HA	1H	3.872
L10	HG	1H	1.384
L10	HN	1H	8.010
L10	N	15N	124.017
L10	QB	1H	1.391
L10	QQD	1H	0.695

R11	HA	1H	3.700
R11	HB2	1H	1.846
R11	HB3	1H	1.690
R11	HE	1H	7.449
R11	HH11	1H	7.174
R11	HH21	1H	7.104
R11	HH22	1H	6.868
R11	HN	1H	8.034
R11	N	15N	117.968
R11	NE	15N	106.522
R11	QD	1H	3.073
R11	QG	1H	1.417
K12	HA	1H	4.072
K12	HD2	1H	1.520
K12	HD3	1H	1.438
K12	HN	1H	7.766
K12	N	15N	117.936
K12	QB	1H	1.839
K12	QE	1H	2.940
K12	QG	1H	1.673
G13	HA2	1H	3.965

G13	HA3	1H	3.630
G13	HN	1H	8.132
G13	N	15N	106.172
F14	HA	1H	4.241
F14	HB2	1H	2.953
F14	HB3	1H	3.153
F14	HN	1H	8.110
F14	HZ	1H	6.773
F14	N	15N	117.161
F14	QD	1H	6.846
F14	QE	1H	7.181
Y15	HA	1H	4.246
Y15	HB2	1H	3.158
Y15	HB3	1H	2.965
Y15	HN	1H	8.202
Y15	N	15N	120.673
Y15	QD	1H	6.771
D16	HA	1H	4.620
D16	HB2	1H	3.112
D16	HB3	1H	2.906
D16	HN	1H	8.341

D16	N	15N	117.348
G17	HN	1H	7.916
G17	N	15N	107.893
G17	QA	1H	3.974
E18	HA	1H	4.064
E18	HN	1H	8.280
E18	N	15N	120.303
E18	QB	1H	2.022
E18	QG	1H	2.349
A19	HA	1H	4.050
A19	HN	1H	8.289
A19	N	15N	121.292
A19	QB	1H	1.319
G20	HA2	1H	3.889
G20	HA3	1H	3.746
G20	HN	1H	8.151
G20	N	15N	104.693
R21	HA	1H	4.155
R21	HB2	1H	1.818

R21	HB3	1H	1.711
R21	HE	1H	7.345
R21	HN	1H	7.723
R21	N	15N	119.079
R21	NE	15N	106.276
R21	QD	1H	3.148
R21	QG	1H	1.572
A22	HA	1H	4.178
A22	HN	1H	8.006
A22	N	15N	121.750
A22	QB	1H	1.427
I23	HA	1H	4.091
I23	HB	1H	1.932
I23	HD1	1H	0.851
I23	HG2	1H	0.913
I23	HN	1H	7.668
I23	N	15N	114.719
I23	QG1	1H	1.246
R24	HA	1H	4.324
R24	HB2	1H	1.902
R24	HB3	1H	1.749
R24	HG2	1H	1.693

R24	HG3	1H	1.558
R24	HN	1H	7.801
R24	N	15N	121.161
R24	QD	1H	3.136
R25	HA	1H	4.269
R25	HB2	1H	1.921
R25	HB3	1H	1.809
R25	HE	1H	7.440
R25	HN	1H	8.170
R25	N	15N	121.077
R25	NE	15N	106.473
R25	QD	1H	3.189
R25	QG	1H	1.675

5.5 Upper distance Limits list

Peak number (#peak) above 10.000 indicates that the constraint is derived from a 2D NOESY spectrum. The rest of the peaks are from a 3D NOESY-HSQC spectrum. CYANA nomenclature.

3	TRP	H	3	TRP	QB	2.92	#peak 3	#SUP 0.95 #QF 0.95
2	ALA	H	3	TRP	H	3.61	#peak 4	#SUP 0.99 #QF 0.99
2	ALA	QB	3	TRP	H	3.51	#peak 6	#SUP 0.99 #QF 0.99
17	GLY	H	18	GLU	H	2.70	#peak 7	#SUP 0.72 #QF 0.72

12 LYS H 13 GLY H 3.92 #peak 8 #SUP 1.00

12 LYS QB 13 GLY H 3.99 #peak 11 #SUP 0.99 #QF 0.99

12 LYS QG 13 GLY H 3.26 #peak 13 #SUP 0.90 #QF 0.90

10 LEU HA 13 GLY H 4.55 #peak 15 #SUP 0.87 #QF 0.87

20 GLY H 21 ARG H 3.43 #peak 18 #SUP 0.99 #QF 0.99

19 ALA H 20 GLY H 3.59 #peak 19 #SUP 0.97 #QF 0.97

18 GLU HA 20 GLY H 3.73 #peak 22 #SUP 1.00 #QF 0.83

19 ALA QB 20 GLY H 3.27 #peak 23 #SUP 0.97 #QF 0.97

18 GLU H 18 GLU QG 3.48 #peak 24 #SUP 0.91 #QF 0.91

18 GLU H 18 GLU QB 3.15 #peak 25 #SUP 0.94 #QF 0.94

18 GLU QG 19 ALA H 4.52 #peak 26 #SUP 1.00

18 GLU QB 19 ALA H 3.64 #peak 27 #SUP 0.98 #QF 0.98

19 ALA H 19 ALA QB 2.88 #peak 28 #SUP 1.00

14 PHE HA 17 GLY H 5.50 #peak 32 #SUP 1.00 #QF 0.96

15 TYR HA 17 GLY H 5.50 #peak 32 #SUP 1.00 #QF 0.96

16 ASP HB2 17 GLY H 5.13 #peak 33 #SUP 1.00

16 ASP HB3 17 GLY H 4.47 #peak 35 #SUP 0.98 #QF 0.98

13 GLY H 14 PHE HA 5.50 #peak 37 #SUP 0.92 #QF 0.92

17 GLY QA 18 GLU H 3.05 #peak 39 #SUP 1.00

17 GLY QA 19 ALA H 3.81 #peak 41 #SUP 0.82 #QF 0.82

19 ALA H 20 GLY HA3 5.10 #peak 42 #SUP 0.99 #QF 0.99

19 ALA H 20 GLY HA2 5.50 #peak 43 #SUP 0.95 #QF 0.95

10 LEU H 10 LEU QQD 3.72 #peak 44 #SUP 0.99 #QF 0.99

20 GLY H 23 ILE H 5.03 #peak 52 #SUP 1.00 #QF 0.93

22 ALA H 23 ILE H 3.36 #peak 53 #SUP 0.93 #QF 0.93

22 ALA H 22 ALA QB 2.83 #peak 55 #SUP 0.96 #QF 0.96

20 GLY HA3 23 ILE H 5.50 #peak 58 #SUP 0.98 #QF 0.98

21 ARG HB3 23 ILE H 5.50 #peak 61 #SUP 0.92 #QF 0.92

23 ILE H 25 ARG QG 5.50 #peak 62 #SUP 0.92 #QF 0.23

22 ALA QB 23 ILE H 3.31 #peak 63 #SUP 0.91 #QF 0.91

23 ILE HB 24 ARG H 3.65 #peak 64 #SUP 1.00 #QF 0.75

24 ARG H 24 ARG HG3 3.75 #peak 66 #SUP 0.98 #QF 0.98

24 ARG H 24 ARG HG2 2.85 #peak 67 #SUP 0.54 #QF 0.54

24 ARG H 25 ARG H 2.76 #peak 68 #SUP 0.81 #QF 0.81

22 ALA H 24 ARG H 5.48 #peak 69 #SUP 0.93 #QF 0.93

23 ILE HA 24 ARG H 3.50 #peak 72 #SUP 0.99 #QF 0.99

23 ILE QG1 24 ARG H 4.40 #peak 74 #SUP 0.98 #QF 0.98

23 ILE QG2 24 ARG H 3.97 #peak 75 #SUP 1.00

23 ILE H 23 ILE QG1 3.54 #peak 77 #SUP 0.99 #QF 0.99

23 ILE H 23 ILE QG2 3.87 #peak 78 #SUP 1.00

18 GLU H 22 ALA H 5.50 #peak 79 #SUP 1.00

21 ARG H 22 ALA H 5.50 #peak 81 #SUP 0.98 #QF 0.98

1 GLY QA 3 TRP H 4.25 #peak 82 #SUP 0.98 #QF 0.98

21 ARG HA 24 ARG H 4.69 #peak 91 #SUP 0.97 #QF 0.84

22 ALA HA 24 ARG H 5.11 #peak 91 #SUP 0.97 #QF 0.84

23 ILE HA 25 ARG H 4.37 #peak 93 #SUP 0.98 #QF 0.98

23 ILE H 25 ARG H 4.35 #peak 94 #SUP 0.91 #QF 0.91

14 PHE H 18 GLU H 5.50 #peak 96 #SUP 0.99 #QF 0.58

18 GLU H 20 GLY H 5.50 #peak 96 #SUP 0.99 #QF 0.58

13 GLY H 14 PHE H 3.36 #peak 97 #SUP 0.65 #QF 0.65

1 GLY QA 2 ALA H 2.80 #peak 101 #SUP 0.99 #QF 0.99

10 LEU QQD 11 ARG H 4.37 #peak 104 #SUP 1.00

7 TRP HA 9 SER H 3.75 #peak 105 #SUP 0.85 #QF 0.85

9 SER HA 12 LYS H 4.41 #peak 106 #SUP 0.93 #QF 0.93

12 LYS H 13 GLY HA2 4.87 #peak 107 #SUP 0.95 #QF 0.95

9 SER H 10 LEU HA 5.50 #peak 108 #SUP 0.95 #QF 0.95

10 LEU HA 12 LYS H 4.49 #peak 109 #SUP 0.94 #QF 0.94

18 GLU H 21 ARG H 5.50 #peak 112 #SUP 1.00

21 ARG H 21 ARG HB3 2.88 #peak 113 #SUP 0.91 #QF 0.91

21 ARG H 21 ARG HB2 3.31 #peak 114 #SUP 0.97 #QF 0.97

21 ARG H 21 ARG QG 4.16 #peak 115 #SUP 1.00

25 ARG H 25 ARG QG 2.59 #peak 119 #SUP 0.83 #QF 0.83

14 PHE H 15 TYR H 5.50 #peak 120 #SUP 0.28 #QF 0.28

22 ALA H 25 ARG H 5.50 #peak 121 #SUP 0.95 #QF 0.20

25 ARG H 25 ARG HE 5.22 #peak 122 #SUP 0.98 #QF 0.98

24 ARG H 24 ARG QD 3.78 #peak 125 #SUP 0.78 #QF 0.78

22 ALA HA 25 ARG H 3.92 #peak 126 #SUP 0.73 #QF 0.73

25 ARG H 25 ARG QD 4.49 #peak 127 #SUP 1.00

23 ILE QG2 25 ARG H 5.06 #peak 128 #SUP 0.99 #QF 0.99

3 TRP HA 6 PHE H 3.95 #peak 129 #SUP 1.00 #QF 0.41

6 PHE H 6 PHE HB2 2.90 #peak 133 #SUP 0.80 #QF 0.80

6 PHE H 6 PHE HB3 2.90 #peak 134 #SUP 0.88 #QF 0.88

11 ARG H 14 PHE H 5.50 #peak 136 #SUP 0.78 #QF 0.78

23 ILE H 23 ILE QD1 4.03 #peak 137 #SUP 0.98 #QF 0.98

8 SER QB 9 SER H 5.40 #peak 139 #SUP 1.00 #QF 0.99

9 SER H 11 ARG HB2 5.00 #peak 140 #SUP 0.85 #QF 0.85

9 SER H 12 LYS QG 5.50 #peak 141 #SUP 0.90 #QF 0.90

9 SER H 10 LEU QB 5.50 #peak 143 #SUP 0.99 #QF 0.99

9 SER H 11 ARG H 5.50 #peak 144 #SUP 0.97 #QF 0.97

11 ARG H 12 LYS H 5.50 #peak 145 #SUP 1.00

7 TRP HA 11 ARG H 5.06 #peak 146 #SUP 0.96 #QF 0.96

9 SER HA 11 ARG H 5.00 #peak 147 #SUP 0.79 #QF 0.79

11 ARG H 11 ARG HB3 4.16 #peak 151 #SUP 1.00

10 LEU QB 11 ARG H 3.05 #peak 152 #SUP 0.69 #QF 0.69

18 GLU HA 21 ARG H 3.47 #peak 153 #SUP 0.85 #QF 0.85

20 GLY HA3 21 ARG H 3.01 #peak 154 #SUP 0.88 #QF 0.88

21 ARG H 24 ARG QD 4.27 #peak 155 #SUP 0.66 #QF 0.66

21 ARG H 22 ALA QB 4.45 #peak 156 #SUP 0.83 #QF 0.83

19 ALA QB 21 ARG H 4.69 #peak 157 #SUP 0.99 #QF 0.99

3 TRP H 4 LYS QB 5.39 #peak 158 #SUP 0.99 #QF 0.99

14 PHE HA 18 GLU H 3.65 #peak 159 #SUP 0.50 #QF 0.50

18 GLU H 19 ALA QB 4.47 #peak 160 #SUP 0.98 #QF 0.98

19 ALA H 22 ALA H 5.50 #peak 161 #SUP 1.00

17 GLY H 19 ALA H 5.24 #peak 162 #SUP 0.97 #QF 0.97

19 ALA H 21 ARG H 4.59 #peak 163 #SUP 0.99 #QF 0.99

14 PHE HA 19 ALA H 3.91 #peak 164 #SUP 0.73 #QF 0.73

18 GLU HA 22 ALA H 3.80 #peak 165 #SUP 0.82 #QF 0.82

20 GLY HA3 22 ALA H 4.95 #peak 166 #SUP 0.99 #QF 0.99

18 GLU QB 22 ALA H 4.19 #peak 168 #SUP 0.87 #QF 0.87

21 ARG HB3 22 ALA H 3.20 #peak 170 #SUP 0.94 #QF 0.94

21 ARG HB2 22 ALA H 4.06 #peak 171 #SUP 0.98 #QF 0.98

21 ARG QG 22 ALA H 3.64 #peak 172 #SUP 0.95 #QF 0.95

19 ALA QB 22 ALA H 5.50 #peak 173 #SUP 1.00

22 ALA H 23 ILE QG1 5.50 #peak 174 #SUP 0.88 #QF 0.88

22 ALA H 23 ILE QG2 5.50 #peak 175 #SUP 0.99 #QF 0.99

5 ASN QB 6 PHE H 3.00 #peak 176 #SUP 0.68 #QF 0.68

9 SER H 10 LEU H 5.50 #peak 180 #SUP 1.00

10 LEU H 12 LYS H 4.89 #peak 181 #SUP 0.99 #QF 0.99

7 TRP HA 10 LEU H 4.01 #peak 182 #SUP 0.73 #QF 0.73

9 SER HB2 10 LEU H 4.00 #peak 184 #SUP 0.97 #QF 0.97

16 ASP H 19 ALA H 5.50 #peak 187 #SUP 0.77 #QF 0.77

25 ARG HB2 25 ARG HE 5.50 #peak 190 #SUP 1.00

11 ARG HB2 11 ARG HE 5.50 #peak 191 #SUP 1.00 #QF 0.66

25 ARG HB3 25 ARG HE 5.50 #peak 191 #SUP 1.00 #QF 0.66

11 ARG QG 11 ARG HE 2.98 #peak 193 #SUP 0.87 #QF 0.87

11 ARG H 11 ARG QD 5.50 #peak 195 #SUP 1.00

5 ASN HA 5 ASN HD22 4.64 #peak 200 #SUP 0.99 #QF 0.99

5 ASN QB 5 ASN HD22 3.89 #peak 201 #SUP 1.00

3 TRP H 3 TRP HE3 4.49 #peak 202 #SUP 0.98 #QF 0.98

3 TRP H 3 TRP HD1 4.24 #peak 203 #SUP 0.98 #QF 0.98

20 GLY H 21 ARG HB3 4.01 #peak 204 #SUP 0.97 #QF 0.97

20 GLY H 22 ALA QB 4.77 #peak 205 #SUP 0.97 #QF 0.97

18 GLU QB 20 GLY H 5.23 #peak 206 #SUP 0.97 #QF 0.97

20 GLY H 21 ARG QG 5.50 #peak 207 #SUP 0.96 #QF 0.96

20 GLY H 24 ARG QD 5.50 #peak 209 #SUP 0.90 #QF 0.90

14 PHE H 16 ASP H 5.50 #peak 211 #SUP 0.73 #QF 0.73

15 TYR H 16 ASP H 5.50 #peak 212 #SUP 1.00 #QF 0.79

16 ASP H 17 GLY H 5.50 #peak 213 #SUP 1.00

16 ASP H 18 GLU HA 5.50 #peak 215 #SUP 0.77 #QF 0.77

15 TYR QD 16 ASP H 5.14 #peak 218 #SUP 1.00

20 GLY H 21 ARG HB2 5.00 #peak 219 #SUP 0.93 #QF 0.93

17 GLY H 18 GLU QB 4.86 #peak 221 #SUP 0.91 #QF 0.91

17 GLY H 19 ALA QB 5.12 #peak 222 #SUP 0.98 #QF 0.98

23 ILE H 24 ARG H 5.50 #peak 223 #SUP 1.00

23 ILE H 24 ARG HA 5.50 #peak 225 #SUP 0.98 #QF 0.98

20 GLY HA2 23 ILE H 3.70 #peak 226 #SUP 0.51 #QF 0.51

5 ASN HA 9 SER H 5.50 #peak 227 #SUP 0.86 #QF 0.86

7 TRP HB2 9 SER H 5.50 #peak 228 #SUP 0.97 #QF 0.97

5 ASN QB 9 SER H 5.50 #peak 229 #SUP 0.92 #QF 0.25

9 SER H 12 LYS QE 5.50 #peak 229 #SUP 0.92 #QF 0.25

9 SER H 10 LEU QQD 5.28 #peak 230 #SUP 0.87 #QF 0.87

12 LYS H 12 LYS QE 4.58 #peak 231 #SUP 0.95 #QF 0.95

12 LYS H 12 LYS QB 3.17 #peak 232 #SUP 1.00

12 LYS H 12 LYS QG 2.83 #peak 233 #SUP 0.82 #QF 0.82

12 LYS H 12 LYS HD2 4.51 #peak 234 #SUP 1.00

12 LYS H 12 LYS HD3 4.51 #peak 235 #SUP 0.97 #QF 0.97

21 ARG H 21 ARG HE 4.82 #peak 236 #SUP 0.99 #QF 0.99

21 ARG H 23 ILE HB 4.96 #peak 237 #SUP 0.49 #QF 0.49

18 GLU QB 21 ARG H 5.26 #peak 238 #SUP 0.95 #QF 0.95

18 GLU QG 21 ARG H 5.50 #peak 239 #SUP 0.98 #QF 0.98

21 ARG QD 22 ALA H 5.35 #peak 240 #SUP 0.95 #QF 0.59

22 ALA H 24 ARG QD 5.50 #peak 240 #SUP 0.95 #QF 0.59

24 ARG HG2 25 ARG H 3.53 #peak 242 #SUP 0.67 #QF 0.67

16 ASP H 16 ASP HB2 3.22 #peak 246 #SUP 0.72 #QF 0.33

2 ALA QB 6 PHE H 5.27 #peak 253 #SUP 0.93 #QF 0.93

5 ASN HD21 6 PHE H 4.27 #peak 254 #SUP 0.80 #QF 0.80

6 PHE H 7 TRP HE3 4.67 #peak 255 #SUP 0.58 #QF 0.58

6 PHE H 6 PHE HD2 5.32 #peak 256 #SUP 0.98 #QF 0.97

6 PHE H 11 ARG HH21 5.50 #peak 256 #SUP 0.98 #QF 0.97

6 PHE H 6 PHE HD1 5.09 #peak 257 #SUP 0.94 #QF 0.94

5 ASN HA 5 ASN HD21 4.90 #peak 260 #SUP 0.99 #QF 0.99

12 LYS H 13 GLY HA3 4.87 #peak 263 #SUP 0.99 #QF 0.99

21 ARG HB3 21 ARG HE 2.91 #peak 264 #SUP 0.79 #QF 0.79

21 ARG HB2 21 ARG HE 3.12 #peak 265 #SUP 0.84 #QF 0.84

21 ARG QG 21 ARG HE 4.07 #peak 266 #SUP 0.99 #QF 0.99

21 ARG QD 21 ARG HE 2.55 #peak 267 #SUP 0.98 #QF 0.98

25 ARG HA 25 ARG HE 4.56 #peak 268 #SUP 0.90 #QF 0.90

3 TRP H 5 ASN H 4.57 #peak 272 #SUP 0.99 #QF 0.99

5 ASN H 5 ASN QB 2.55 #peak 278 #SUP 0.71 #QF 0.71

4 LYS QB 5 ASN H 2.91 #peak 279 #SUP 0.84 #QF 0.84

2 ALA QB 5 ASN H 4.89 #peak 280 #SUP 0.98 #QF 0.98

4 LYS QD 5 ASN H 5.17 #peak 281 #SUP 1.00

5 ASN H 5 ASN HD21 3.75 #peak 286 #SUP 0.97 #QF 0.24

5 ASN H 5 ASN HD22 4.09 #peak 287 #SUP 0.99 #QF 0.99

2 ALA HA 5 ASN HD21 3.56 #peak 288 #SUP 0.84 #QF 0.84

2 ALA HA 5 ASN HD22 5.50 #peak 289 #SUP 0.99 #QF 0.99

5 ASN QB 5 ASN HD21 2.93 #peak 290 #SUP 0.99 #QF 0.99

25 ARG H 25 ARG HA 2.79 #peak 293 #SUP 0.94 #QF 0.94

4 LYS QB 6 PHE H 4.68 #peak 294 #SUP 0.73 #QF 0.73

15 TYR HA 16 ASP H 3.41 #peak 295 #SUP 0.94 #QF 0.94

3 TRP H 4 LYS QD 5.19 #peak 297 #SUP 1.00

2 ALA HA 5 ASN H 3.70 #peak 299 #SUP 0.79 #QF 0.79

16 ASP H 17 GLY QA 4.55 #peak 301 #SUP 0.26 #QF 0.26

2 ALA H 5 ASN HD21 5.50 #peak 302 #SUP 0.99 #QF 0.99

2 ALA H 5 ASN HD22 5.50 #peak 303 #SUP 0.96 #QF 0.96

2 ALA QB 5 ASN HD21 4.46 #peak 304 #SUP 0.97 #QF 0.97

3 TRP H 6 PHE H 5.05 #peak 305 #SUP 0.90 #QF 0.90

4 LYS HA 6 PHE H 4.20 #peak 306 #SUP 0.55 #QF 0.55

11 ARG H 11 ARG HE 3.68 #peak 309 #SUP 0.88 #QF 0.88

9 SER H 11 ARG HE 5.39 #peak 310 #SUP 0.89 #QF 0.89

11 ARG H 11 ARG HH21 5.04 #peak 313 #SUP 0.95 #QF 0.95

9 SER H 11 ARG HH21 4.96 #peak 315 #SUP 0.97 #QF 0.97

9 SER H 11 ARG HH22 5.50 #peak 316 #SUP 0.92 #QF 0.92

10 LEU H 11 ARG HE 5.50 #peak 318 #SUP 0.99 #QF 0.99

10 LEU H 11 ARG HH21 5.50 #peak 319 #SUP 0.95 #QF 0.95

10 LEU H 11 ARG HB3 5.50 #peak 320 #SUP 0.90 #QF 0.59

10 LEU H 12 LYS QG 5.50 #peak 320 #SUP 0.90 #QF 0.59

7 TRP HB3 11 ARG H 4.53 #peak 321 #SUP 0.36 #QF 0.36

9 SER H 12 LYS H 5.50 #peak 323 #SUP 0.98 #QF 0.98

11 ARG HE 12 LYS H 5.50 #peak 324 #SUP 0.92 #QF 0.92

10 LEU QQD 12 LYS H 5.42 #peak 326 #SUP 0.97 #QF 0.97

9 SER HA 13 GLY H 4.90 #peak 327 #SUP 0.82 #QF 0.82

13 GLY H 16 ASP HB2 5.50 #peak 328 #SUP 0.28 #QF 0.28

10 LEU QQD 13 GLY H 4.57 #peak 329 #SUP 0.59 #QF 0.59

15 TYR QD 17 GLY H 5.50 #peak 330 #SUP 0.70 #QF 0.70

17 GLY H 18 GLU QG 5.06 #peak 331 #SUP 0.99 #QF 0.99

16 ASP HA 18 GLU H 4.68 #peak 332 #SUP 0.95 #QF 0.95

18 GLU H 21 ARG HB3 4.58 #peak 334 #SUP 0.57 #QF 0.57

14 PHE QB 19 ALA H 4.47 #peak 335 #SUP 0.76 #QF 0.76

20 GLY H 22 ALA H 3.37 #peak 337 #SUP 0.97 #QF 0.97

20 GLY H 23 ILE QG2 5.50 #peak 338 #SUP 0.74 #QF 0.74

21 ARG H 24 ARG H 5.50 #peak 339 #SUP 0.98 #QF 0.98

11 ARG QD 11 ARG HE 2.80 #peak 340 #SUP 0.99 #QF 0.99

18 GLU HA 21 ARG HE 5.07 #peak 341 #SUP 0.64 #QF 0.64

11 ARG HA 11 ARG HE 4.51 #peak 342 #SUP 0.99 #QF 0.99

23 ILE H 23 ILE HB 2.68 #peak 349 #SUP 0.78 #QF 0.78

20 GLY HA3 24 ARG H 5.50 #peak 350 #SUP 0.91 #QF 0.91

20 GLY HA2 24 ARG H 5.35 #peak 351 #SUP 0.91 #QF 0.91

2 ALA H 3 TRP QB 4.99 #peak 352 #SUP 0.92 #QF 0.92

10 LEU H 10 LEU QB 2.42 #peak 353 #SUP 0.65 #QF 0.65

3 TRP H 3 TRP HE1 4.89 #peak 354 #SUP 0.96 #QF 0.96

7 TRP H 7 TRP HE3 3.43 #peak 357 #SUP 0.67 #QF 0.67

7 TRP H 7 TRP HB2 2.55 #peak 359 #SUP 0.51 #QF 0.51

3 TRP H 4 LYS H 2.81 #peak 360 #SUP 0.89 #QF 0.89

2 ALA H 4 LYS H 4.31 #peak 361 #SUP 0.95 #QF 0.95

7 TRP H 11 ARG HH21 3.99 #peak 363 #SUP 0.90 #QF 0.90

7 TRP H 11 ARG HH22 3.81 #peak 364 #SUP 0.75 #QF 0.75

3 TRP HA 4 LYS H 3.50 #peak 365 #SUP 0.98 #QF 0.98

1 GLY QA 4 LYS H 3.99 #peak 368 #SUP 0.79 #QF 0.79

4 LYS H 4 LYS QE 5.15 #peak 370 #SUP 0.61 #QF 0.61

4 LYS H 4 LYS QB 2.72 #peak 374 #SUP 0.91 #QF 0.91

4 LYS H 4 LYS QD 3.67 #peak 375 #SUP 1.00

4 LYS H 4 LYS HG3 3.54 #peak 376 #SUP 0.98 #QF 0.98

4 LYS H 4 LYS HG2 3.54 #peak 377 #SUP 0.99 #QF 0.99

8 SER H 9 SER H 5.50 #peak 381 #SUP 1.00

8 SER H 10 LEU H 5.50 #peak 382 #SUP 1.00

8 SER H 11 ARG H 5.50 #peak 382 #SUP 1.00

8 SER H 11 ARG HE 4.14 #peak 383 #SUP 0.96 #QF 0.96

8 SER H 8 SER HA 2.85 #peak 384 #SUP 0.97 #QF 0.97

7 TRP HA 8 SER H 3.45 #peak 386 #SUP 0.99 #QF 0.99

7 TRP HB2 8 SER H 3.32 #peak 388 #SUP 0.74 #QF 0.74

8 SER H 8 SER QB 3.21 #peak 389 #SUP 0.96 #QF 0.96

8 SER H 10 LEU QB 5.50 #peak 390 #SUP 0.77 #QF 0.77

15 TYR H 15 TYR HA 2.92 #peak 393 #SUP 0.99 #QF 0.99

15 TYR H 15 TYR HB2 2.90 #peak 394 #SUP 0.89 #QF 0.89

15 TYR H 15 TYR HB3 2.90 #peak 395 #SUP 0.76 #QF 0.76

14 PHE QD 15 TYR H 3.28 #peak 398 #SUP 0.49 #QF 0.49

14 PHE H 14 PHE QD 3.31 #peak 401 #SUP 0.75 #QF 0.75

14 PHE H 14 PHE QE 4.37 #peak 402 #SUP 0.95 #QF 0.95

12 LYS H 14 PHE H 5.50 #peak 405 #SUP 0.99 #QF 0.99

15 TYR H 17 GLY H 5.50 #peak 406 #SUP 0.70 #QF 0.70

3 TRP HD1 4 LYS H 5.50 #peak 407 #SUP 0.99 #QF 0.97

7 TRP H 7 TRP HD1 5.50 #peak 407 #SUP 0.99 #QF 0.97

4 LYS H 5 ASN HD21 5.50 #peak 408 #SUP 0.90 #QF 0.35

7 TRP H 8 SER HA 5.50 #peak 409 #SUP 0.97 #QF 0.97

8 SER H 11 ARG HH22 4.17 #peak 410 #SUP 0.82 #QF 0.82

8 SER H 11 ARG HH21 3.91 #peak 412 #SUP 0.96 #QF 0.96

9 SER HB3 10 LEU H 4.00 #peak 413 #SUP 0.90 #QF 0.90

8 SER HA 11 ARG H 3.40 #peak 414 #SUP 0.42 #QF 0.42

23 ILE H 24 ARG HG2 4.25 #peak 416 #SUP 0.50 #QF 0.50

3 TRP H 4 LYS QG 5.34 #peak 298

4 LYS H 4 LYS QG 3.04 #peak 376

4 LYS QG 5 ASN H 4.19 #peak 282

6 PHE H 6 PHE QB 2.50 #peak 133

9 SER H 9 SER QB 2.58 #peak 49

9 SER H 12 LYS QD 5.34 #peak 143

12 LYS H 12 LYS QD 3.63 #peak 235

12 LYS H 13 GLY QA 4.12 #peak 107

12 LYS QD 13 GLY H 3.72 #peak 12

13 GLY QA 16 ASP H 4.84 #peak 216

15 TYR H 15 TYR QB 2.35 #peak 395

24 ARG H 24 ARG QB 3.35 #peak 65

25 ARG H 25 ARG QB 2.52 #peak 118

17 GLY QA 21 ARG HA 6.38 #peak 10062

24 ARG HA 24 ARG HB2 3.00 #peak 10100

24 ARG HA 25 ARG H 3.00 #peak 10109

24 ARG HA 24 ARG HB3 3.00 #peak 10111

21 ARG HA 21 ARG HB2 3.00 #peak 10133

21 ARG HA 21 ARG HB3 3.00 #peak 10134

23 ILE HA 23 ILE HB 3.00 #peak 10138

17 GLY QA 18 GLU QB 6.13 #peak 10278

21 ARG HA 24 ARG QD 4.03 #peak 10125

21 ARG HA 21 ARG QG 3.88 #peak 10128

21 ARG HA 21 ARG HE 4.35 #peak 10129

12 LYS HA 12 LYS QG 3.88 #peak 10140

3 TRP QB 3 TRP HD1 3.88 #peak 10020

3 TRP QB 3 TRP HE3 3.88 #peak 10335

12 LYS QB 12 LYS QE 7.06 #peak 10186

15 TYR HA 15 TYR QE 5.33 #peak 10245

11 ARG HA 15 TYR QE 7.16 #peak 10247

15 TYR H 15 TYR QE 6.31 #peak 10288

15 TYR H 15 TYR QD 5.28 #peak 10316

15 TYR HA 15 TYR QD 5.38 #peak 10317

12 LYS HA 15 TYR QD 6.70 #peak 10318

10 LEU HA 15 TYR QD 7.28 #peak 10319

2 ALA H 2 ALA QB 4.02 #peak 10077

22 ALA QB 25 ARG QG 7.40 #peak 10079

22 ALA QB 23 ILE QG2 7.41 #peak 10082

22 ALA QB 23 ILE QD1 7.54 #peak 10083

23 ILE HB 23 ILE QD1 4.02 #peak 10088

23 ILE QG2 24 ARG HA 5.02 #peak 10114

23 ILE HA 23 ILE QG2 4.02 #peak 10144

23 ILE HA 23 ILE QD1 4.02 #peak 10145

10 LEU HA 10 LEU QQD 5.10 #peak 10158

18 GLU QG 22 ALA QB 6.93 #peak 10199

19 ALA HA 22 ALA QB 4.02 #peak 10203

2 ALA QB 3 TRP HZ3 4.89 #peak 10250

10 LEU QQD 14 PHE QD 9.73 #peak 10251

2 ALA QB 3 TRP HE3 4.02 #peak 10343

9 SER QB 10 LEU H 3.12 #peak 10173

11 ARG QG 11 ARG QH1 4.01 #peak 10039

11 ARG QH2 15 TYR QD 4.99 #peak 10232

17 GLY QA 20 GLY QA 6.16 #peak 10061

20 GLY QA 21 ARG H 2.74 #peak 10160

21 ARG QB 21 ARG HE 2.97 #peak 10331

21 ARG QB 22 ALA QB 6.29 #peak 10081