Thesis Submitted for the Degree of *Candidata Scientiarum*

# The Three-Dimensional NMR-Structure of the Membrane-Permeabilizing Pediocin-Like Antimicrobial Peptide Curvacin A in Lipid Micelles

By Helén Sophie Haugen


*Department of Molecular Biosciences*

Faculty of Mathematics and Natural Sciences
UNIVERSITY OF OSLO

# Publication resulted from this study

Haugen H.S., Fimland G., Nissen-Meyer J., Kristiansen P.E

**Three-Dimensional Structure in Lipid Micelles of the Pediocin-like Antimicrobial Peptide Curvacin A.**

Submitted to Biochemistry

# Aknowledgements

Nuclear magnetic measurements (NMR) were carried out at the Swedish NMR Centre, University of Gothenburg. Special thanks to Göran B. Karlsson for letting me work at the Hasselblad Laboratory.

NMR sample preparation, resonance assignments and structure calculations were carried out at the Department of Molecular Biosciences, University of Oslo.

I would like to thank Dr. Per Eugen Kristiansen for introducing me to the field of NMR. You have been very patient with me and always shared your knowledge, even when I had a hard time of it! I have enjoyed our trips to Gothenburg and sharing office with you.

Furthermore, I would like to thank Professor Jon Nissen-Meyer for taking me on as a student. Your enthusiasm and interest in my project has been remarkable. It is very inspiring for a student to have a professor that takes so much part in one's project. I have enjoyed my years here!

Also, I wish to extend a warm thank-you to Dr. Gunnar Fimland. Your help and guidance have been irreplaceable, and you are a constant source of knowledge and great ideas!

To my fellow-students and members of the pediocin-group: Ane, Mariam, Grete, Line, Kari, Per. You have made my days fun, even when every experiment went down the drain…

Last, to the people who mean the most to me; my family and friends. Thank you for always being there for me, for believing in me and for giving me the privilege of your care and love!

Oslo, August 2005

Helén Haugen

# Abstract

Antimicrobial peptides produced by lactic acid bacteria (LAB) are extensively studied, but little is know about their three-dimensional structures. Solution structures of these peptides will be useful to gain information about their mode of action, which again is important for further use of these substances in industry and medicine. Curvacin A is an antimicrobial peptide that belongs to the family of pediocin-like bacteriocins.

Curvacin A was produced from its natural LAB producer, *Lactobacillus curvatus* LTH1174, and purified by ion-exchange- and reverse phase chromatography. Circular Dichroism-experiments revealed that curvacin A was unstructured in water, but became structured upon interactions with membrane-mimicking environments such as dodecyl phosphocholin (DPC)-micelles.

The three-dimensional structure of curvacin A in DPC-micelles has been elucidated by the use of nuclear magnetic resonance (NMR)-spectroscopy. Curvacin A was shown to contain three regions: an N-terminal S-shaped β-sheet like domain (residues 2-15), a central polar helix (residues 19-24) and an amphiphilic C-terminal helix (residues 29-39). The C-terminal tail consists of only two residues (G40 and M41) and seems to be unstructured. There was a hinge between these regions, enabling the regions to move relative to each other.

# Table of content

# 1 Introduction

## 1.1 Antimicrobial peptides produced by lactic acid bacteria

Historically, fermentation of food and beverage is an ancient tradition where the intrinsic properties of lactic acid bacteria (LAB) are utilized. Fermentation is a way of preserving food from harmful degradation by bacteria and other food-borne pathogens [1]. The protective power of LABs has been ascribed to the presence of their main end product, lactate [2]. Lactate creates an acidic environment that effectively prevents growth of almost all potential spoilage microorganisms [2, 3].

In recent years, the focus of attention has moved to antimicrobial peptides (AMPs) secreted by LABs that kill other similar bacterial strains. Specifically, ribosomally synthesized membrane-permeabilizing AMPs, termed bacteriocins, produced by a wide variety of LABs have been of interest. These peptides are usually cationic, small (< 50-60 residues) and contain a putative amphiphilic or hydrophobic region [4].

In search of possible active agents to be used in areas such as pharmaceutical industry and agriculture, biologically naturally occurring substances that might prove effective against different pathogens are of interest. As the flux of antibiotic-resistant pathogens increases, it has become paramount to find new sources of future effective drugs and food-preservatives to replace soon-to-be ineffective antibiotics [5]. Bacteriocins show promising features in this respect [3, 6, 7]. Many are produced naturally in fermented food, and have been implemented in inhibiting numerous food spoilage organisms [4, 8].

## 1.2 Classification of bacteriocins produced by LAB:

Peptide bacteriocins produced by LAB have been classified into two groups based on their primary structure [9] (Figure 1.2.1).

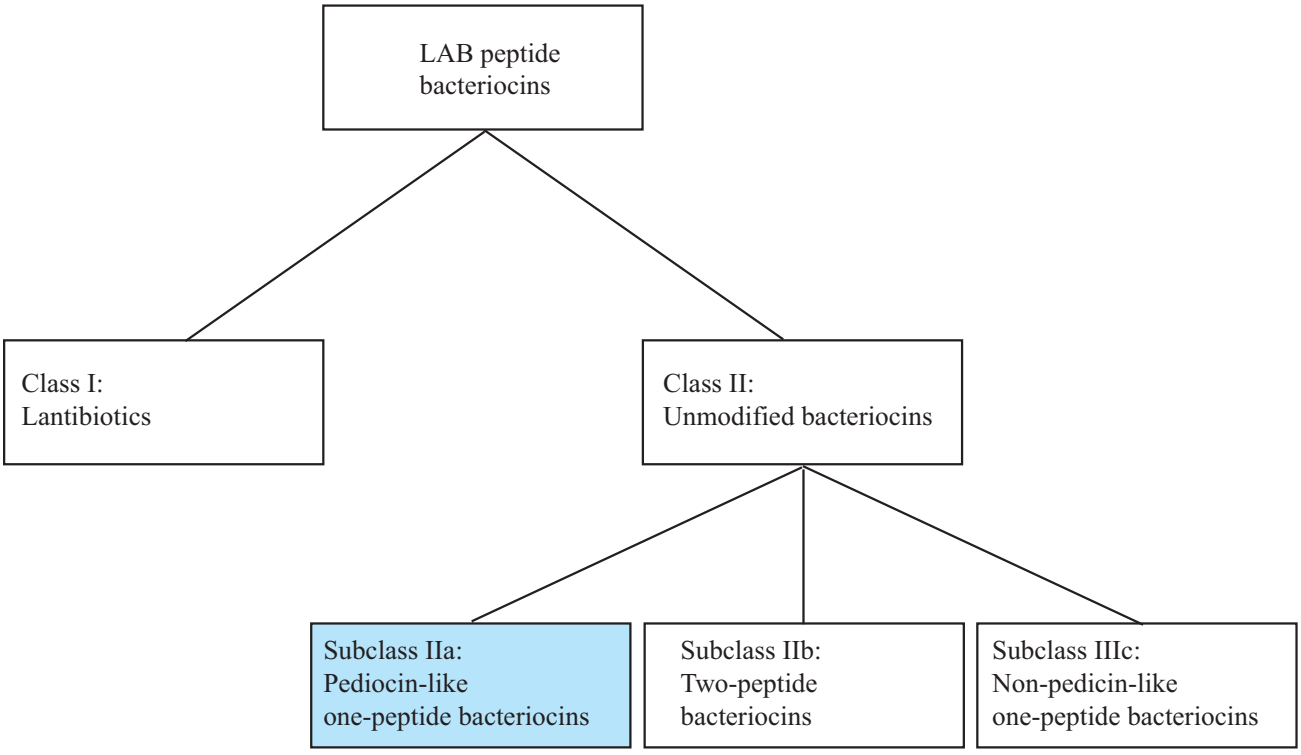


Figure 1.2.1: Classification of peptide bacteriocins produced by LAB. The LAB peptide bacteriocins are divided into two groups, class I and class II. Class II is further divided into three subgroups.

The lantibiotics (class I) are post-translationally modified peptides. They all contain the modified residues lanthionine and/or β-methyllanthionine [4, 10, 11]. In addition, other modified residues such as D-alanine and dehydrated amino acids may be present [12, 13].

The unmodified peptides in class II are further divided into three subclasses [3]. The listeria active subclass IIa bacteriocins kill cells by permeabilizing the cell membrane [14]. They are often termed pediocin-like bacteriocins due to their resemblance to pediocin PA-1 (one of the first AMPs in this subclass to be identified and characterized), and this term is used here [15]. The two-peptide bacteriocins (subclass IIb) consist of two different peptides. Both peptides must be present for the bacteriocin to display antimicrobial activity [16]. Subclass IIc consists of a diverse group of non-pediocin-like one-peptide bacteriocins [16].

## 1.3 Biosynthesis and secretion of bacteriocins:

Production of bacteriocins requires at least four genes [17-22]. These genes are found in close proximity, usually organized in one or two operons. One gene encodes the peptide precursor (prepeptides), a second the immunity protein that renders the bacteria immune to the bacteriocins it produces. A third gene encodes a membrane-associated ABC transporter: The prepeptides are transported across the membrane concomitantly with cleavage of a leader sequence. Finally, a fourth gene encodes an accessory protein necessary for secretion of the bacteriocin. Figure 1.3.1 schematically illustrates biosynthesis and secretion of bacteriocins.

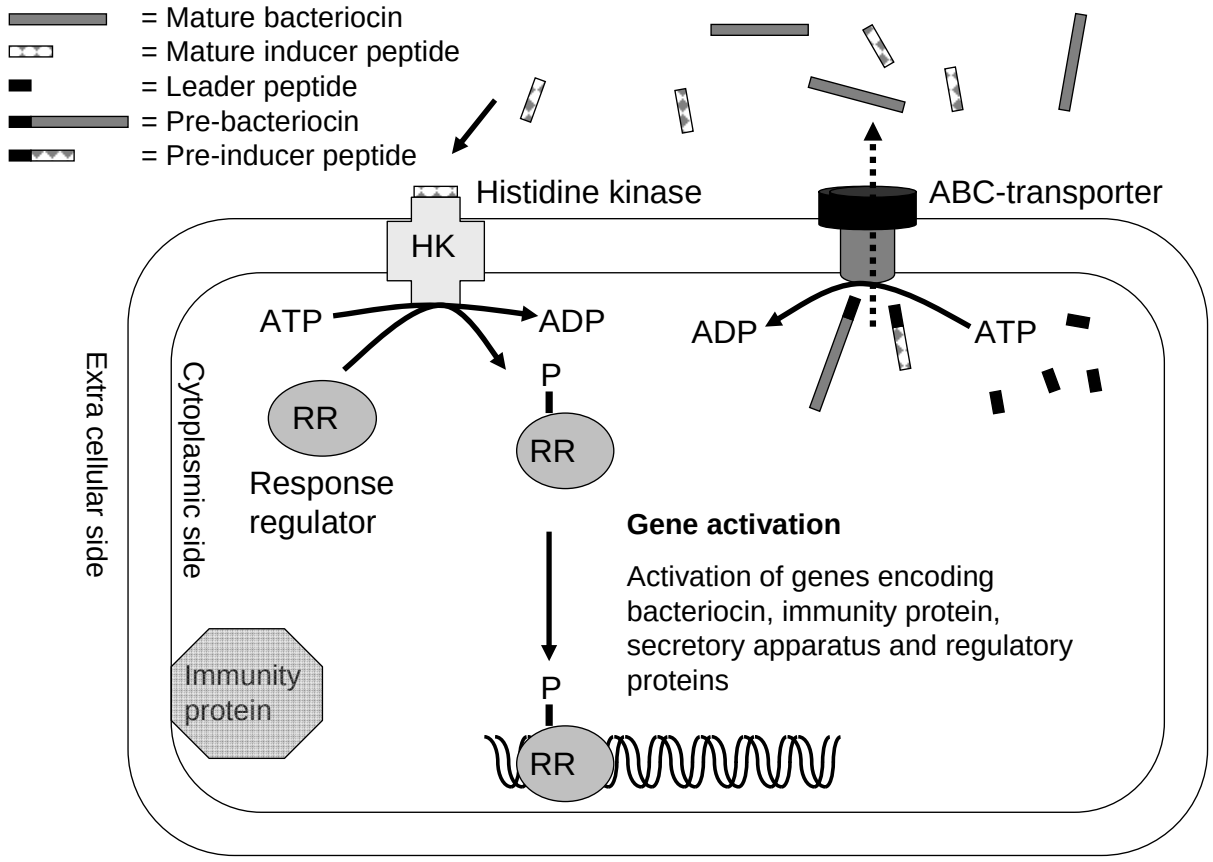


Figure 1.3.1 Illustration of biosynthesis and secretion of bacteriocins. A histidine kinase is activated by mature inducer peptide, and phosphorylates a response regulator. This response regulator binds to the DNA, and the genes indicated in the figure are activated. Pre-bacteriocin and pre-inducer peptide are transported across the membrane concomitantly with cleavage of a leader sequence.

The prepeptide contains an N-terminal leader sequence, which is cleaved off C-terminally of a double glycine motif. The 15 to 30 residues long leader sequence presumably facilitates interactions with the transporter and keeps the bacteriocin inactive until it is secreted from the cell [19-24].

Some bacteriocins (e.g curvacin A, sakacin P, carnobacteriocin B2 and enterocin A) are produced under regulation of a three-component signal transduction system, encoded by three additional genes [25, 26].

## 1.4 Group IIa bacteriocins:

Peptide chains of the pediocin-like bacteriocins may roughly be divided into two regions based on their primary structure (Figure 1.4.1) [27, 28]. The N-terminal region contains the highly conserved Y-G-N-G-V/L, or “pediocin-box”, motif. The region is hydrophilic and cationic, and contains two cysteine residues linked by a disulfide bridge. The C-terminal region is less conserved and is hydrophilic or amphiphilic.

Based on multiple sequence alignment of the C-terminal part, pediocin-like bacteriocins have been divided into three groups (Figure1.4.1).

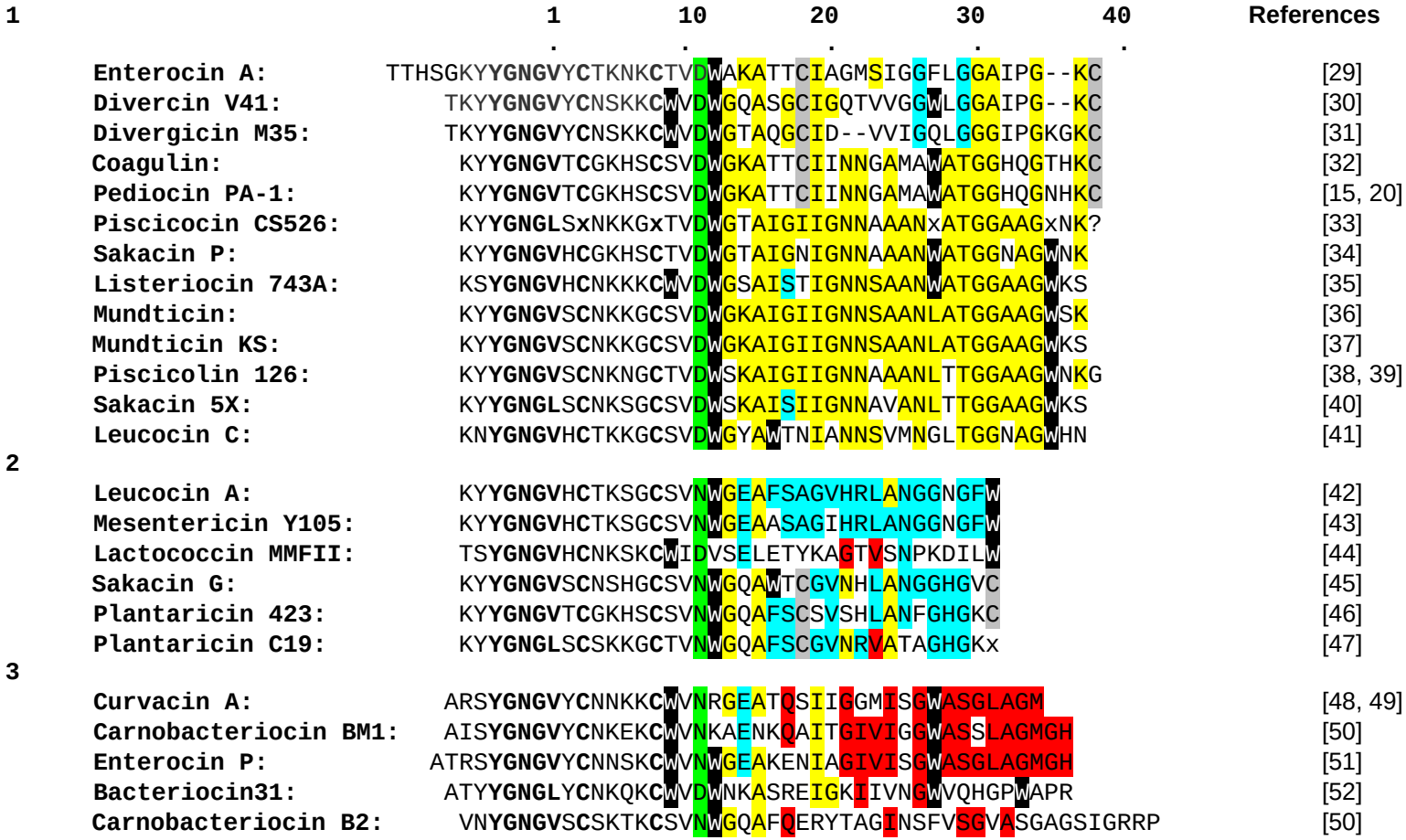


Figure 1.4.1: Multiple sequence alignment of pediocin-like subgroup IIa bacteriocins highlighting the YGNGV/L "pediocin box" motif (bold face) and conserved cysteine residues (bold face) in the N-terminal half, tryptophan residues (black boxes), and conserved residues in the C-terminal half of the peptides (yellow, blue, gray, green and red boxes). The C-terminal half is more diverse than the N-terminal half, and the classification of the peptides into three groups (indicated by 1, 2 and 3 in the figure) is thus based on sequence similarities and differences in the C-terminal half of the peptides. There is a flexible hinge at the conserved Asp17 (green boxes) in peptides of group 1, and presumably also at Asn17/Asp17 (green boxes) in peptides of group 2 and 3. This hinge separates the b-sheet N-terminal domain and the hairpin-like C-terminal domain. Note that in numbering the residues (as indicated above in the sequence), residue number 2 before the well conserved YGNGV motif is in all cases referred to as residue 1, since this residue is the first residue in most -but not all – of the peptides. The PILEUP program of the Genetics Computer Group sequence analysis program package (Wisconsin Package Version 8.1, Genetic Computer Group) was used to set up the alignments [27]

### 1.4.1 Three-dimensional structure of the pediocin-like bacteriocins

NMR-structures of four different pediocin-like bacteriocins have been reported so far; leucocin A (group 2; Figure 1.4.1) [53], carnobacteriocin B2 (group 3; Figure 1.4.1) [54], sakacin P (group 1; Figure 1.4.1) [16] and a modified version of sakacin P where an additional disulfide bridge has been introduced [16].

As all the peptides are unstructured in water, but become structured when introduced to a membrane-mimicking environment [16, 43, 53-55], they have to be dissolved in a structure-inducing solvent in order for their structure to be analyzed. The peptides must be dissolved in a membrane-mimicking environment such as dodecyl phosphocholin (DPC) micelles or trifluoroethanol (TFE) to provide the right conditions for structuring of the peptides. TFE is known to induce helical structuring in proteins, possibly at the expense of β-sheet structures [56], and thus structures calculated from TFE-dissolved peptides might show an unrealistically high content of α-helix. Even so, peptides can benefit from being dissolved in TFE as this stabilizes secondary-structure formation at the cost of hydrogen bonds between water and the peptide [56].

Leucocin A was the first pediocin-like bacteriocin whose structure was elucidated [53]. The peptide, which belongs to group 2 (Figure 1.4.1), was shown to contain an antiparallel three-stranded β-sheet-like structure from residue 2-16 and this was followed by an amphiphilic α-helix spanning residues 17-31, and what appeared to be an unstructured C-terminal tail from residue 32 to 37.

The group 1 pediocin-like bacteriocin sakacin P (Figure 1.4.1) and the mutated variant sakacin P [N24C+44C] were subsequently also shown to contain an N-terminal three-stranded β-sheet-like structure (from residue 1 to 17) followed by an amphiphilic α-helix between residue 18 and 33 [16]. Moreover, it was shown that the C-terminal tail (residues 34-43 for sakacin P and 34-44 for sakacin P [N24C+44C]) folds back on the amphiphilic helix, thus forming a hairpin-like structure (from residue 18 to 43/44), which is stabilized by a disulfide bridge (in sakacin P [N24C+44C] ) or by two tryptophan residues (in sakacin P) that position themselves in the membrane interface (see Figure 1.4.2.1) [16, 57]. It was also shown that this hairpin-like structure and the N-terminal β-sheet-like structure were joined by a hinge (residue 17) and one thus obtains two domains (the N-terminal β-

sheet-like structure and the C-terminal hairpin-like structure) that may move relative to each other (Figure 1.4.2.1) [16].

This probably describes the structure of all the pediocin-like bacteriocins that belong to group 1 and 2 (Figure 1.4.1) as they have a well conserved primary structure in the N-terminal half and they all contain either the hairpin-stabilizing disulfide bond between a C-terminal cysteine residue and a cysteine residue in the middle of the peptide, or a hairpin-stabilizing tryptophan residue near the C-terminal end (Figures 1.4.1 and 1.4.2.2).

It is however unclear whether or not the pediocin-like bacteriocins that belong to group 3 (Figure 1.4.1) also form the three-dimensional structure described for group 1 and 2 peptides, since the group 3 peptides all lack the cysteine residues in the C-terminal half (and thus the hairpin-stabilizing disulfide bond) and the tryptophan residue near the C-terminal end. The NMR structure of one group 3 peptide, carnobacteriocin B2, has been reported, and the structure revealed only an amphiphilic α-helix from residue 18 to 39 [54]. There was no apparent well-defined structuring in the N-terminal half nor the C-terminal tail.

### 1.4.2 Mode of action and orientation in the membranes of pediocin-like bacteriocins

Hybrid bacteriocins constructed by joining the N-terminal domain of one peptide with the C-terminal domain of another (using the hinge as the recombination point) display target cell specificity similar to the bacteriocin from which the C-terminal end was derived [58]. This indicates that the hairpin membrane-permeabilizing C-terminal domain is the major specificity determinant in pediocin-like bacteriocins.

Alterations in the C-terminal region by site directed *in vitro* mutagenesis often alters the target cell specificity of the peptide, which further supports the conclusion that the C-terminal domain is the specificity domain for the pediocin-like bacteriocins [57, 59]. Apparently, interactions with lipids and/or proteins in the interphase and/or hydrophobic phase of the cell membrane determine to a large extent the specificity [58]. Recent studies implicate part of the mannose PTS permease as a possible receptor molecule for the pediocin-like bacteriocins [60-65].

The cationic N-terminal β-sheet-like domain has through site directed *in vitro* mutagenesis and peptide-binding studies been pointed out as the domain that binds the bacteriocin to the target cell surface [66, 67]. This initial interaction between the target cell and the β-sheet-like domain is electrostatic and presumably non-specific. The hydrophobic/amphiphilic C-terminal hairpin domain penetrates then into the hydrophobic part of the target-cell membrane [12, 28, 57, 68], the hinge between the two domains providing the structural flexibility, which enables the C-terminal hairpin domain to dip into the membrane. The penetration causes leakage through the membrane, which in turn may cause dissipation of the proton motive force (PMF). The cells try to restore this loss of PMF by ATP consumption, which again leads to depletion of the cell's ATP pool [36, 69-71]. In Figure 1.4.2.1, the structure and orientation in the target-cell membrane of pediocin-like bacteriocins and its interaction with an immunity protein is depicted.

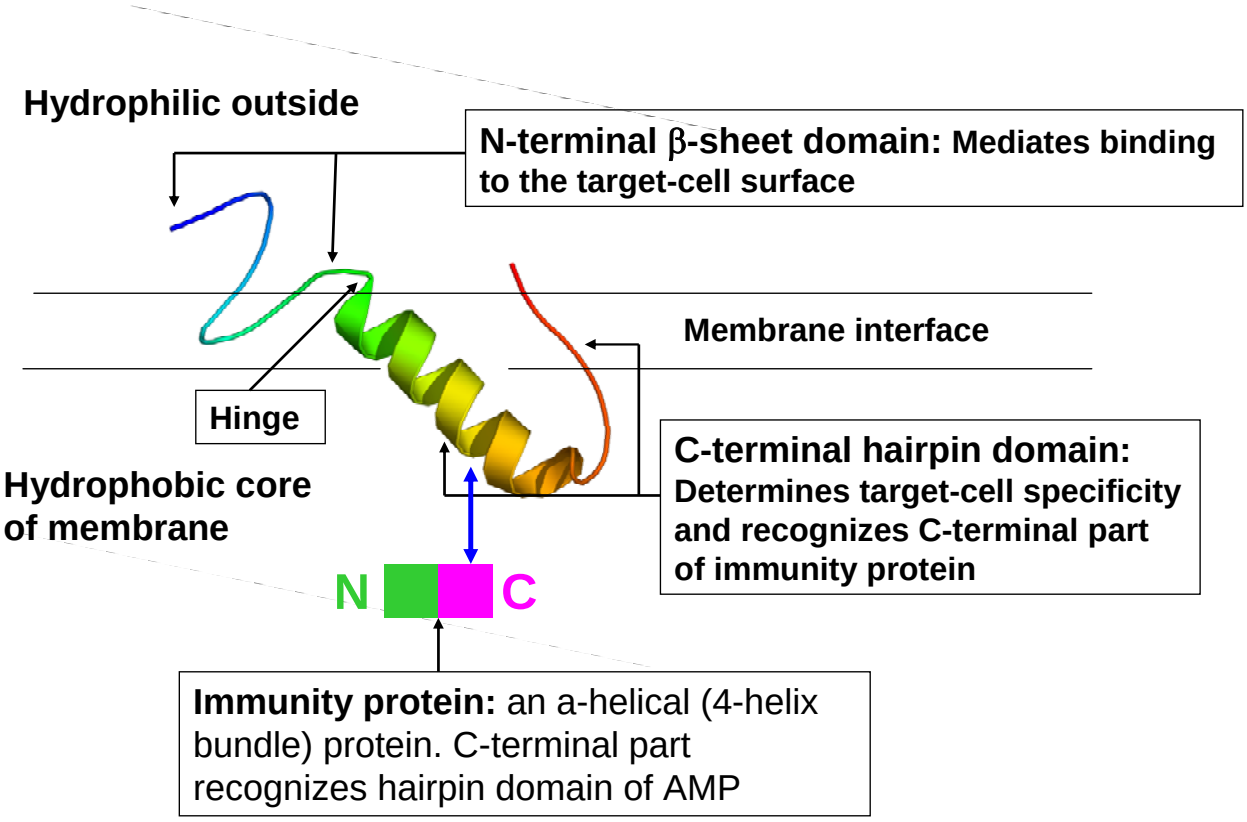


Figure 1.4.2.1 A cartoon depiction of the structure and orientation in target-cell membrane of a pediocin-like bacteriocin, and interaction with an immunity protein. The bacteriocin is divided into two discrete domains, separated by a hinge: the N-terminal β-sheet-like domain that mediates binding to the target-cell surface and the C-terminal hairpin domain that determines target-cell specificity and interacts with the immunity protein.

The very similar primary structure of the N-terminal domains (Figure1.4.1) is presumably the reason why target cells do not seem to discriminate between different pediocin-like bacteriocins upon initial interaction between the peptide and target cells. However, mutational analysis reveals that charged residues in the N-terminal half also may influence target-cell specificity to some extent [67].

All pediocin-like bacteriocins have a central tryptophan residue, usually at position 18 (Figure 1.4.1). This well-conserved tryptophan residue positions itself in the water-membrane interface (as is common for tryptophan residues in membrane penetrating polypeptides [72]), and aids in positioning the C-terminal domain correctly in the membrane (Figure1.4.2.2) [27].

In addition and as discussed above, most of the pediocin-like bacteriocins (all those that belong to groups 1 and 2; Figure 1.4.1) have either a disulfide bridge or tryptophan residue C-terminally. Both these features stabilize the hairpin structure [57]. The disulfide bond does so by covalently binding the C-terminal end to the central part of the peptide and the tryptophan residue by positioning itself in the water-membrane interface in the same manner as the central tryptophan residue (Figure 1.4.2.2).

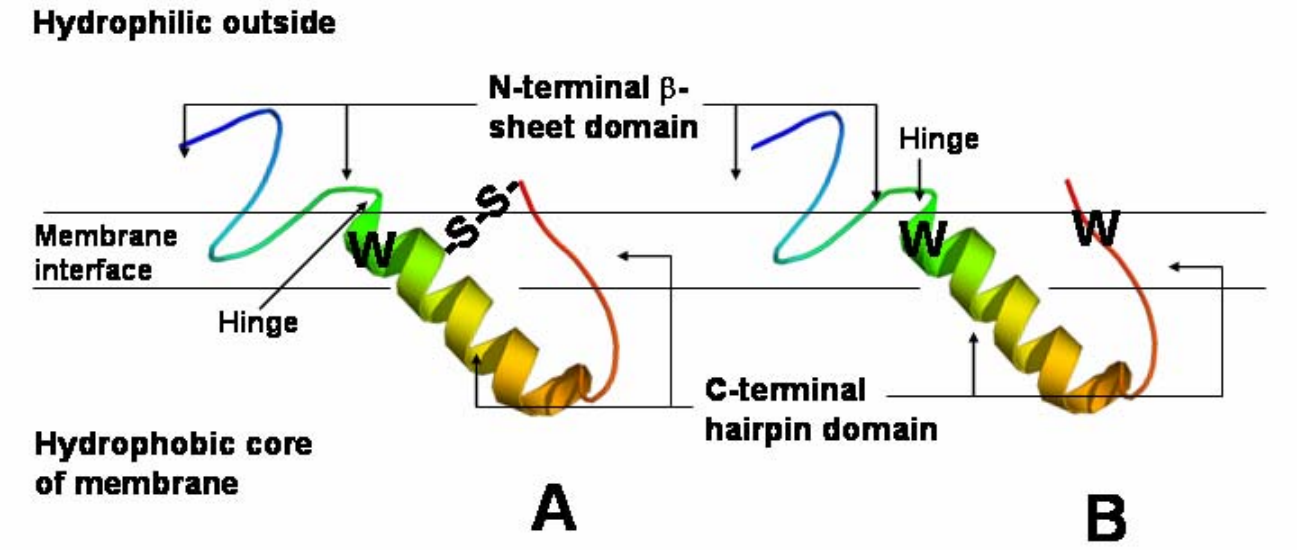


Figure 1.4.2.2: A cartoon depiction of the structure and orientation in membranes of pediocin-like bacteriocins. A pediocin-like bacteriocin in which the C-terminal hairpin structure is stabilized by (A) a disulfide bridge and (B) interface-localized tryptophan residues near the C-terminal end and in the middle of the bacteriocin. Tryptophan residues that become localized in the membrane-water interface are indicated by W and the disulfide bridge by –S-S- [27].

## 1.5 Aim of this study

As discussed above, there is no additional disulfide bridge in group 3 pediocin-like bacteriocins, and the last tryptophan residue is positioned more towards the centre of the C-terminal domain (Figure 1.4.1). The above model describing the structure of pediocin-like bacteriocins that belong to group 1 and 2 and how these bacteriocins position themselves in membranes (Figures 1.4.2.1 and 1.4.2.2) may consecuently not hold for the bacteriocins that belong to group 3. Carnobacteriocin B2 is the only group 3 pediocin-like bacteriocin whose three-dimensional structure has been analyzed, and the central amphiphilic α-helix was the only structural unit which was identified [54]. The aim of this study was to determine the three-dimensional structure of the group 3 pediocin-like bacteriocin curvacin A, and thereby gain more insight into the three-dimensional structure of the pediocin-like bacteriocins that belong to this subgroup.

It was of interest to see if these peptides also form a similar two-domain structure as the peptides that belong to group 1 and 2. Moreover, insight into the three-dimensional structure of group 3 peptides may reveal how these peptides interact with target membranes, information which is vital for understanding at a molecular level their mode of action. This information may in turn be important for possible future commercial use of these peptides in food industry and medicine.

Three-dimensional structures of polypeptides may be obtained by both X-ray crystallography and Nuclear Magnetic Resonance (NMR)-spectroscopy. NMR-spectroscopy was chosen in this study for determining the three-dimensional structure of curvacin A.

Curvacin A is a small molecule (4307 Da) that is unstructured in water and requires a membrane to be structured. A membrane-mimicking environment is created by using trifluoroethanol (TFE) or micelle-forming lipids such as as dodecyl phosphocholin (DPC). In such an environment, NMR experiments are easily performed whereas it is not feasible to obtain the crystals needed for X-ray crystallography under these conditions.

Two different procedures to produce sufficient amounts of curvacin A for NMR-analysis were investigated. One was to clone the curvacin A-enconding gene into an *E. coli* over-expression

system and use this system for production and isolation of curvacin A. The other was to isolate curvacin A from the natural curvacin A-producing strain, *Lactobacillus curvatus* LTH1174.

# 2 Brief Description of Important Methods used in this Study:

## 2.1 Brief introduction to NMR-theory: concepts and applications

In 1946, Edward Mills Purcell [73] and Felix Bloch [74] independently demonstrated the NMR-phenomena: absorption of radio-frequency energy due to energy transitions corresponding to different states of nuclear spins in a strong, constant applied magnetic field [73]. In NMR, as the name indicates, it is the nucleus that is of interest. Nuclei are sensitive towards the electronic environment they find themselves in, and as a result nuclei resonate at different (or shifted) frequencies depending on the local electronic environment surrounding the atom. Structural components and functional groups can be identified based on the observed resonances [75].

An intrinsic property of many nuclei is the spin angular momentum, which gives rise to magnetic moments [76]. Angular momentum is quantized [77].

Electrons, $^{1}H$- $^{13}C$- and $^{15}N$-nuclei share a common spin quantum number s=1/2 and nuclear spin quantum number $I_s = \pm 1/2$. Most biological NMR experiments utilize one or more of the above nuclei for investigation. The following discussion considers the nuclei with s=1/2 but is equally applicable for considering the case of electrons.

There are two possible states for an s=1/2 nuclei: spin up (denoted $\uparrow$ or $\alpha$) and spin down (denoted $\downarrow$ or $\beta$) associated I= +1/2 and -1/2, respectively [76]. When no magnetic field is present, states with different values of quantum number I are degenerate, i.e. their energy is equal. However, when an external applied field is present the different states are no longer degenerate. As a result, the energy levels are split in two when a field is present [77].

The difference in energy is given by

$$\Delta E = E\alpha - E\beta = h\nu = \gamma \hbar B_0 I,$$

where Eα and Eβ denote the energy for the α- and β-states, γ is the gyromagnetic ratio (a measure of how strongly magnetic a nuclide is [78]), ħ is the Planck constant h divided by 2π, $B_0$ is the steady magnetic field applied and I is the quantum magnetic number.

The application of a magnetic field to a charged particle like the nucleus causes it to precess about the direction of the field [77]. The rate at which the vectors precess depends on the nucleus type and is proportional to the applied magnetic field [76, 77]. At equilibrium the frequency of precession is called the Larmor frequency, $\nu_L$, and takes the value

$$\nu_L=-(\gamma B_0)/2\pi.$$

At equilibrium the population distribution between the two states follows Boltzmanns distribution [77]:

$$N\alpha/N\beta= e(-\Delta E/kT),$$

where Nα and Nβ are the populations of the two states, k=Boltzmanns constant and T is the temperature.

At equilibrium Nα is slightly more populated than Nβ. This excess of nuclear spins gives rise to a bulk magnetisation Mz along the +z-axis (defined axis for equilibrium net magnetisation). In the xy-plane, however, all the vectors are randomly distributed and their individual contributions cancel out, i.e. there is no net magnetisation in this plane [77-79].

Nuclear magnetic resonance occurs when the nucleus changes its spin state driven by the absorption of a quantum of energy [78]: By applying radio frequency (RF)-pulses to a system, transitions between energy states can be induced. Traditionally, this is pictured as the net magnetisation vector along the z-axis flipping down onto the xy-plane [78, 79]. Also, due to phase coherence, net magnetisation in the xy-plane can now be observed [77, 79]. Phase coherence is the alignment of all magnetisation vectors in the xy-plane: Their contributions no longer cancel out. The evolution of magnetisation (which is a product of decay and resonance frequency) is measured [77].

There are essentially two relaxation pathways in which the system returns to the original state that is measured in an NMR-experiment [75-77]: 1)Relaxation from the xy-plane to the z-axis and 2) relaxation within the xy-plane.

In time, the Mz component relaxes back from the xy-plane to the +z-axis [78, 79]. This is referred to $T_1$-relaxation or longitudinal relaxation time. It is also referred to as spin-lattice relaxation time, because the excess energy is released to the "lattice" (the surrounding system of energy states) when spins revert from the $\beta$ to the $\alpha$ state [76].

Also, the loss of coherence among spins due to the fact that some spin faster and other slower than the actual Larmor frequency, contributes to relaxation. Immediately after a RF-pulse, there is an orderly bunch of spins as mentioned above [76]. In time, the spins fan out so that in the end all the spins are randomly distributed in the xy-plane. This fanning out is due to random fluctuating interactions and field inhomogenities and causes the signal to vanish [75]. The random fluctuating interactions give rise to $T_2$ -relaxation or transverse relaxation time. This is also called spin-spin relaxation time to reflect that it involves the relative orientations of the spins [76].

Through the two relaxation pathways, the system relaxes back to equilibrium at a rate dependant on the time constants $T_1$ and $T_2$. The signal produced is called an FID (Free Induction Decay), and decreases exponentially. A mathematical procedure called a Fourier Transform (FT) shifts the signal from a time domain to frequency domain scale [76] (Fig 2.1.1).

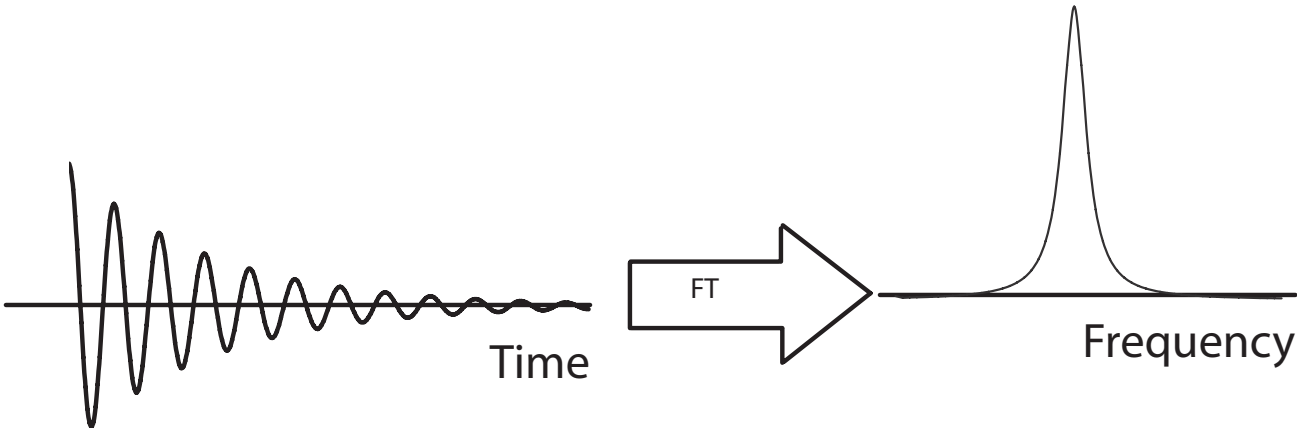


Figure 2.1.1: Schematic description of how the shift from time-domain to frequency-domain by Fourier Transform is achieved for a NMR-spectra.

By applying different pulses in sequences, one can obtain different information. Techniques and pulse sequences have been devised to extract sequential as well as conformational connectivities between for example amino acids. The procedures used in this thesis and the information derived thereof are described later in the text.

### 2.1.1 Information from chemical shifts:

All nuclei resonate at a specific resonance, but the exact value of resonance can vary for even the same type of nuclei, all depending on the local environment. This phenomenon is called chemical shift.

Chemical shifts arise from small changes in electron density around a specific nucleus, and the electronic environment therefore influences the difference in shifts seen for different nuclei. That is, the electrons induce a small additional electronic field either opposing to or adding to the external magnetic field. This leads to shielded and deshielded nuclei, respectively [75]. Hence, a nucleus resonates at slightly different frequencies depending on the surrounding electron densities. Chemical shifts are reported relative to a reference standard to ensure that the values are independent of the applied field and thus a common table of values can be used by all spectroscopists [76].

Standard chemical shift values for nuclei bonded to different neighbours are tabulated and used to identify the signal arising from different functional groups in a spectrum. For example, hydrogen atoms of an amino acid resonate at a recognizable frequency depending on which atom it is bonded to. There is thus a specific pattern of signals from each amino acid, and this is used to identify each residue type in a given protein spectrum. The signals expected for the 20 common amino acids in COSY and/or TOCSY $^{1}$H-$^{1}$H 2D correlation spectrums are shown in Figure 2.1.1.1 below. The chemical shift values correspond to amino acids in random coil structures.

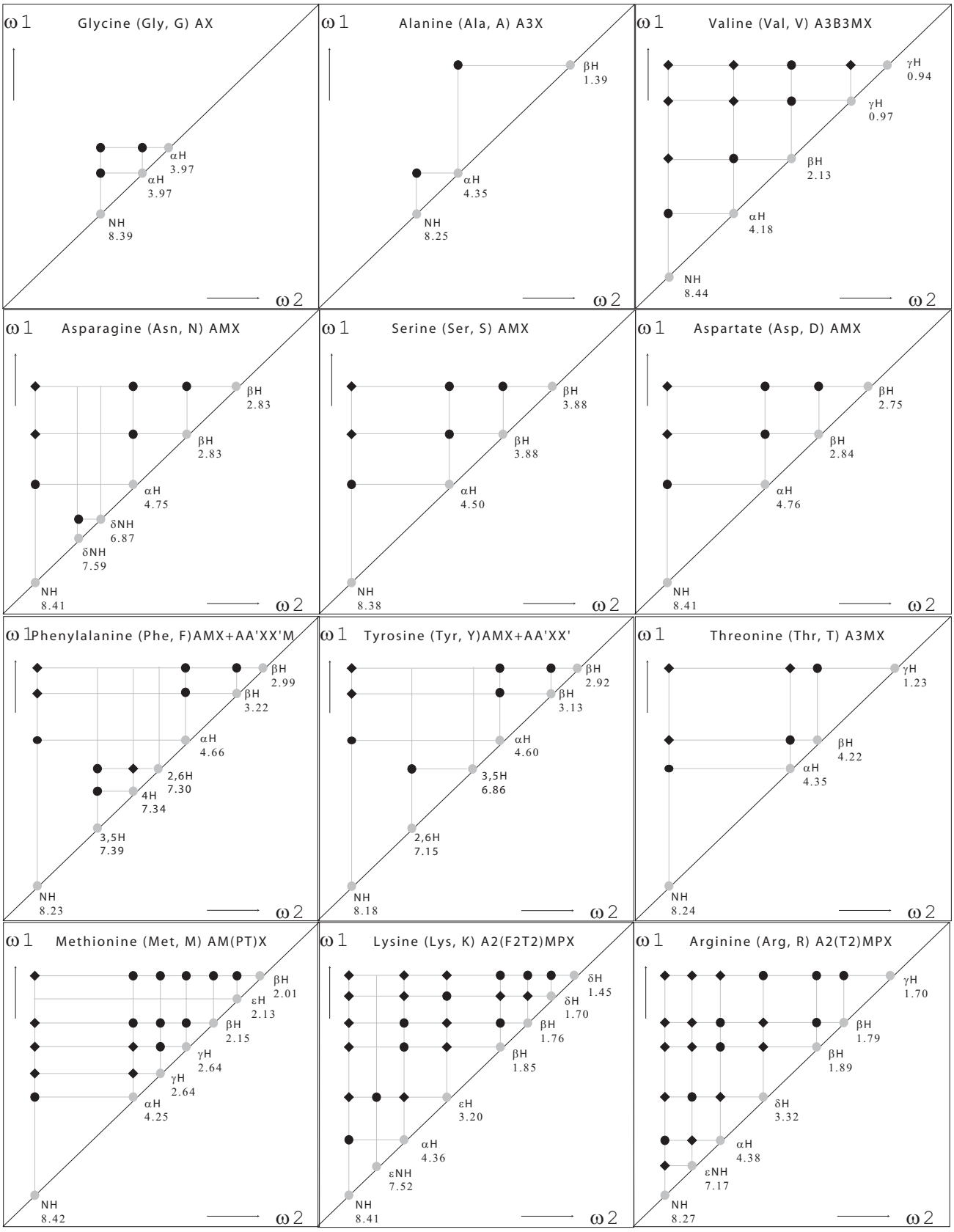
ω1
Glycine (Gly, G) AX
αH 3.97
αH 3.97
NH 8.39
ω2
ω1
Alanine (Ala, A) A3X
βH 1.39
αH 4.35
NH 8.25
ω2
ω1
Valine (Val, V) A3B3MX
γH 0.94
γH 0.97
βH 2.13
αH 4.18
NH 8.44
ω2
ω1
Asparagine (Asn, N) AMX
βH 2.83
βH 2.83
αH 4.75
δNH 6.87
δNH 7.59
NH 8.41
ω2
ω1
Serine (Ser, S) AMX
βH 3.88
βH 3.88
αH 4.50
NH 8.38
ω2
ω1
Aspartate (Asp, D) AMX
βH 2.75
βH 2.84
αH 4.76
NH 8.41
ω2
ω1
Phenylalanine (Phe, F)AMX+AA'XX'M
βH 2.99
βH 3.22
αH 4.66
2,6H 7.30
4H 7.34
3,5H 7.39
NH 8.23
ω2
ω1
Tyrosine (Tyr, Y)AMX+AA'XX'
βH 2.92
βH 3.13
αH 4.60
3,5H 6.86
2,6H 7.15
NH 8.18
ω2
ω1
Threonine (Thr, T) A3MX
γH 1.23
βH 4.22
αH 4.35
NH 8.24
ω2
ω1
Methionine (Met, M) AM(PT)X
βH 2.01
εH 2.13
βH 2.15
γH 2.64
γH 2.64
αH 4.25
NH 8.42
ω2
ω1
Lysine (Lys, K) A2(F2T2)MPX
δH 1.45
δH 1.70
βH 1.76
βH 1.85
εH 3.20
αH 4.36
εNH 7.52
NH 8.41
ω2
ω1
Arginine (Arg, R) A2(T2)MPX
γH 1.70
βH 1.79
βH 1.89
δH 3.32
αH 4.38
εNH 7.17
NH 8.27
ω2

Figure 2.1.1.1

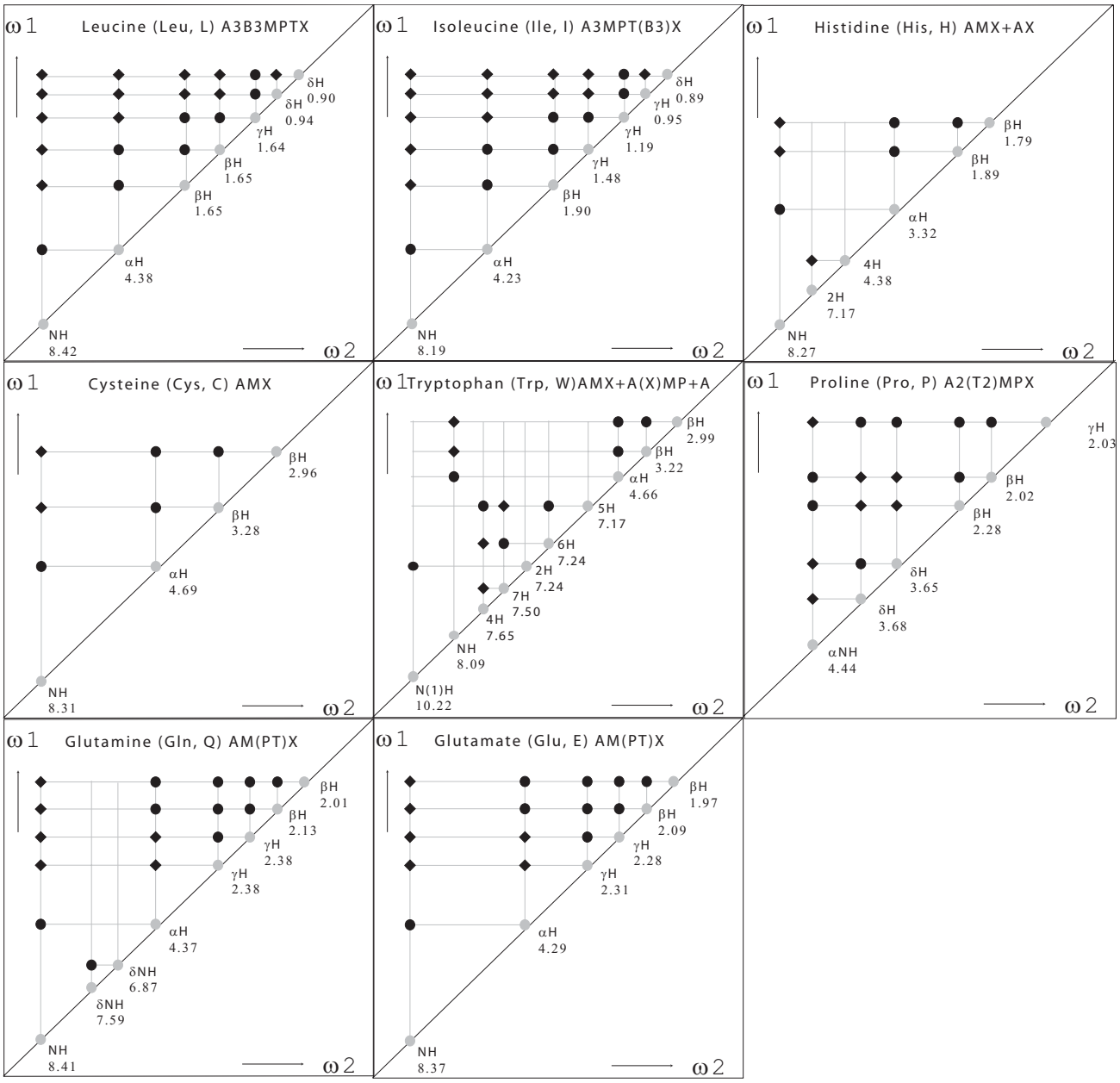


Figure 2.1.1.1 continues. An illustration showing the spin systems of the 20 common amino acid residues in a COSY/TOCSY $^{1}$H-2D NMR experiment. Only peaks on the left hand side of the diagonal are shown. In a real spectrum, identical peaks reflected about the diagonal can be visible. A characteristic pattern of signals results from each amino acid from which the amino acid can be identified. All peaks may be shown in a TOCSY spectrum. Filled gray circles show diagonal peaks, filled black circles show cross peaks visible in a COSY spectrum, filled black squares show additional cross peaks visible in a TOCSY spectrum [80]. Random coil $^{1}$H chemical shift values are included [80].

The spin systems have rather complicated names including letters and numbers. The letters refer to the different coupled spins. The numbers indicate the number of magnetically equivalent protons. For a thorough description, the reader is referred to "NMR of proteins and nucleic acids", chapter 5 by Wüthrich [80]. In the work done in this thesis, only the specific patterns of the spin systems and the labelling of the atoms are of importance.

Figure 2.1.1.2 exemplifies how the atoms in an amino acid are labelled in NMR-theory. The first carbon in the side chain (attached to the carbonyl carbon) is called the α-carbon. The attached proton(s) is(are) the α-proton(s). The subsequent side chain atoms are named by following the Greek alphabet. Thus, a proton is labelled α, β, γ, δ or ε according to which carbon in the side chain it is attached to. The proton attached to the backbone amide is labelled HN. The protons on the aromatic side chains are labelled according to standard numbering.

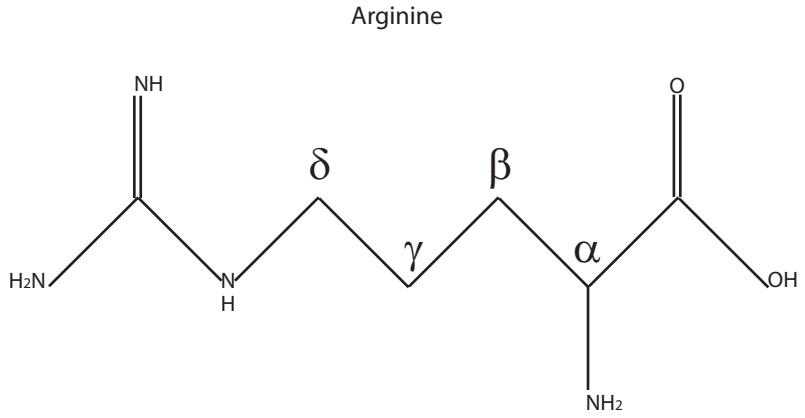


Figure 2.1.1.2 The amino acid arginine with side chain carbons labelled as is standard in NMR

Chemical shifts are among the easiest data collected from a NMR-spectrum. The values are sensitive to the local environment and a deviation from random coil values in chemical shifts is observed when a particular residue is known to be part of a secondary structure [81-84]. $^{15}$N, Hα- and amid-protons experience a net upfield shift when in an α-helical conformation and a net downfield shift when in a β-sheet conformation. α- and carbonyl-$^{13}$C-shifts display the opposite, with downfield shift for α-helical conformations and upfield shift for β-sheet conformations. For flexible or unstructured regions, shifts are close to random coil values.

Given the above trends, one can quantitatively measure a certain part of a structure's flexibility. Chemical shift values close to random coil values of a peptide known to be e.g. an α-helix suggest that this helix is quite flexible, whereas chemical shifts far from random coil values points at a rather rigid structure [83].

Chemical Shift Index (CSI)

One can predict the secondary structure of a given segment by examining the chemical shifts. Especially α-proton chemical shifts show a strong correlation to secondary structural elements. A method devised to explore this relationship is the Chemical Shift Index (CSI) [84].

Backbone angles

There is a relation between chemical shifts of backbone atoms, and secondary structure as mentioned above. In addition, these parameters are linked to backbone angles. By comparing chemical shifts derived from a new compound with those of an already known structure, ϕ– and ψ–angles can be predicted [82, 83, 85]. These angles can then be used as an additional parameter alongside NOEs and J-couplings when calculating the structure. Figure 2.1.1.3 illustrates between which atoms the ϕ– and ψ–angles are found.

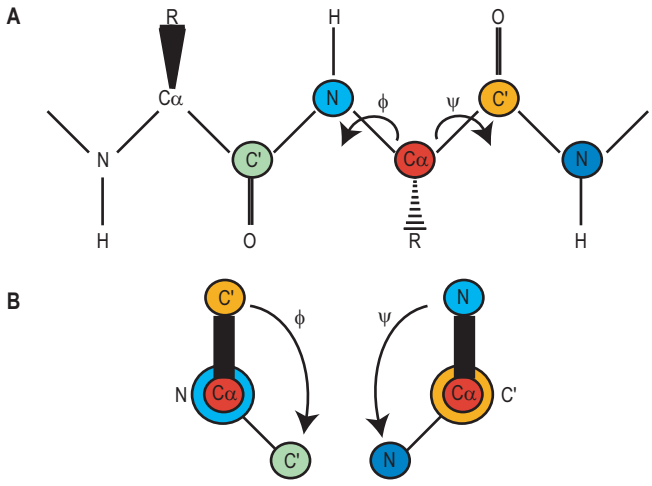


Figure 2.1.1.3 The torsion angles ϕ and ψ between backbone atoms in amino acids. The angle of rotation around the N-Cα bond is the ϕ-angel, and rotation around the Cα –C' bond is the ψ-angel. **A** Plane of a backbone peptide with the side chains oriented above of the plane (filled wedge) or below the plane (striped wedge). The arrows indicate the rotational bonds. **B** Projection of the peptide in part A, viewed along the bond of rotation from the atom drawn as a small circle to the atom drawn as a big circle. The thick black line represents the bond between the atom drawn as a small circle and its neighbouring atom. The thin black line represents the bond between the atom drawn as a big circle and its neighbouring atom. The atoms are coloured as in A.

### 2.1.2 Scalar (spin-spin) couplings

Electrons form chemical bonds between nuclei, and this interaction is known as scalar or spin-spin coupling. The interaction is measured by the scalar coupling constant $^{n}J_{ab}$, where n denotes the number of covalent bonds separating the two nuclei a and b [86].

With scalar coupling (J-coupling), magnetisation is transferred through bonds: All the spins influence each other in a predictable way, and this can aid in determination and labelling of functional groups [75]. In a J-coupled correlation spectrum, all protons within an amino acid have the possibility to couple with each other. (Correlation with protons on adjacent amino acids is so small that it is not visible.) The basis for this is described under the TOCSY section (2.1.5.1.2). The most important couplings from a spectroscopist's point of view is for n=1-4 [86]. Couplings between nuclei cause splitting of the NMR-signal. Weakly and strongly coupled two-spin systems are referred to as AX- and AB-spin systems, respectively [86].

J-values for n=1-3 are dependant not only on internuclear distance but also molecular structure. As such, one can obtain possible parameters for structure calculations from scalar couplings [79]. Particularly $^{3}$J-couplings (vicinal couplings) are useful in this respect. The value is dependant on substituents, the distance between the two carbon atoms involved, H-C-C-H bond angle, along with torsional and/or dihedral angles, with the latter angle being of greatest interest [79].

$^{3}$J-couplings vary in magnitude in a bond-angle dependent manner. The relationship between the coupling constant J and the dihedral angel $\phi$ (Figure 2.1.1.3) are given by the Karplus equation [76]. Consequently, J-couplings can be used to measure dihedral angles and thus conformational arrangements around bonds.

### 2.1.3 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as internal standard; Chemical shift referencing

When listing chemical shifts for the various spin-systems, it is important that there is a common reference. Wishart and colleagues [87] have investigated the properties of different commonly used

reference compounds. The conclusion reached was that DSS is the best standard for both $^{1}H$ and $^{13}C$ chemical shift references. Also, updated tables of random coil values calculated experimentally with DSS as an internal standard are provided [88]. These tabulated values are in good agreement with statistically calculated chemical shifts [89], which further supports the reported work. This study employs DSS as an internal standard.

### 2.1.4 Linewidth and signal:noise-ratio

As the molecular mass increases, so does the rotational correlation time. This has a profound effect on the linewidth of the peaks; the longer the correlation time, the broader the lines [86]. Consequently, the peaks become less well defined and they overlap. The decreased sensitivity can be understood by considering the small $^{3}J$ scalar couplings that coherence transfer relies on: as the linewidths broaden, they become comparable or even greater than the actual J-values we try to measure [86]. Linewidth thus affects the resolution of a spectrum, but it also affects the signal-to-noise ratio [86].

Signal-to-noise ratio, or sensitivity, is proportional to the number of scans in an experiment [90]. More specifically, it improves with the square root of the number of scans in an experiment [77].

### 2.1.5 Two-dimensional NMR-techniques: COSY, TOCSY, NOESY and HSQC.

2D NMR was introduced for the first time by Jean Jeener during a lecture at Ampére Summer School in former Yugoslavia in 1971 [91], and the principles he outlined still govern today's techniques. Four stages of a 2D-experiment can be outlined; preparation, evolution, mixing and acquisition (Figure 2.1.5.1) [86, 90, 92]:

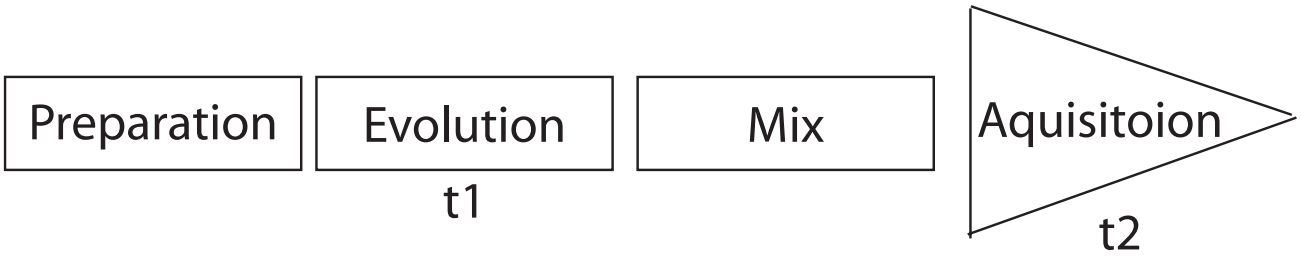

Fig 2.1.5.1: Anatomy of a 2D experiment. Each box represents a stage in the recording process of a 2D NMR spectrum. The two time domains $t_1$ and $t_2$ are acquired during the evolution period and the acquisition period, respectively.

The first stage, the preparation phase, can consist of a single pulse or a more complex sequence of pulses that perturbate the system from its equilibrium state to a desired preparatory state. During the evolution period (stage 2), successive increments of $t_1$ give the F1 dimension in the finished spectrum. In stage three, the mixing, coherence is transferred from one spin to another. This part of the experiment defines what kind of correlation is observed between the two dimensions, i.e. what kind of information the spectrum provides. Thus, here lies the key to obtaining useful information from a 2D spectrum. In the fourth and last stage, the acquisition step, the FID is recorded and this in turn gives the F2 dimension.

The two frequency domains are, as mentioned, derived from t1 and t2, where t1 is incremented during successive runs. The effect is that the evolutions of spins are recorded at different stages and a picture of the chain of events is gained. In other words, we can follow the magnetisation as it travels from one nucleus to another and from this establish which nuclei are connected and in what way they are connected (through bonds, space or both).

The duration of the mixing time determines which peaks that are visible in a given spectrum. With short mixing times, the magnetisation has no time to travel down a side chain. As a result, the α-protons will generally be more intense than the β-protons along the amide shift, which again will be more intense than a γ-proton. Longer mixing time is essential to be able to see the protons at high shifts, most notably the δ- and γ-protons of valine-, isoleucine- and leucine residues.

### 2.1.6 COSY (COrrelation SpectroscopY)

COSY is one of the first 2D techniques to be described [92]. Cross peaks arise from protons separated by two or three bonds, where the coupled spins cause coherence transfer [86]. For proteins, particularly pairs of protons separated by three bonds ($^3J$ scalar coupling), such as methyl groups in the fingerprint region and glycin-residues, are important peak contributors in COSY spectra [86]. A limitation of this technique is that correlations between spins more than three bonds apart cannot easily be assigned.

### 2.1.7 TOCSY (TOtal Correlation SpectroscopY)

Also called HOHAHA (HOmonuclear HArtmann-HAhn), a TOCSY experiment exploits that magnetism can be transferred through several couplings during the mixing time [93]. Pulses are applied such that the Hartmann-Hahn condition is satisfied [77]. Under these conditions magnetisation is transferred very efficiently between coupled nuclei, e.g. an amino acid. Only intraresidual couplings are observable in TOCSY [86]. In a TOCSY experiment one obtains information about which spins are coupled together as a spin system, and hence what type of amino acid that gives rise to a particular signal. By utilizing a long enough mixing time, magnetism is transferred completely through a spin system, and a complete set of cross-peaks for spin-coupled systems can be recorded [77].

Unfortunately, at long mixing times some signals from amino acids with shorter chains can have lost intensity. One should therefore use several different mixing times when recording spectra and compare the results [86].

TOCSY is a widely used method for locating each amino acid residue in a protein with respect to the type of residue, but not to a specific residue. For example, cysteine, histidine, phenylalanine, tryptophan and tyrosine residues all give rise to the same array of connected cross peaks (Figure 2.1.1.1). We can from a TOCSY spectrum only identify a spin system as belonging to this group, but not which amino acid in particular. Alanine, however, has a unique signal and a specific assignment can be made from TOCSY. Even so, if there is more than one alanine residue in the sequence one can only state that a signal comes from an alanine but not which one [94].

<u>Identification of spin-systems; TOCSY</u>
The first step in the assignment procedure is to decide which peaks belong to the same spin-system; that is to the same amino acid. A TOCSY-spectrum provides this kind of information. As mentioned in the description of TOCSY-spectra above, only the type of spin-system can be determined in this step. Figure 2.1.7.1 illustrates how the peaks in a spin-system are connected in different regions of a TOCSY-spectrum.

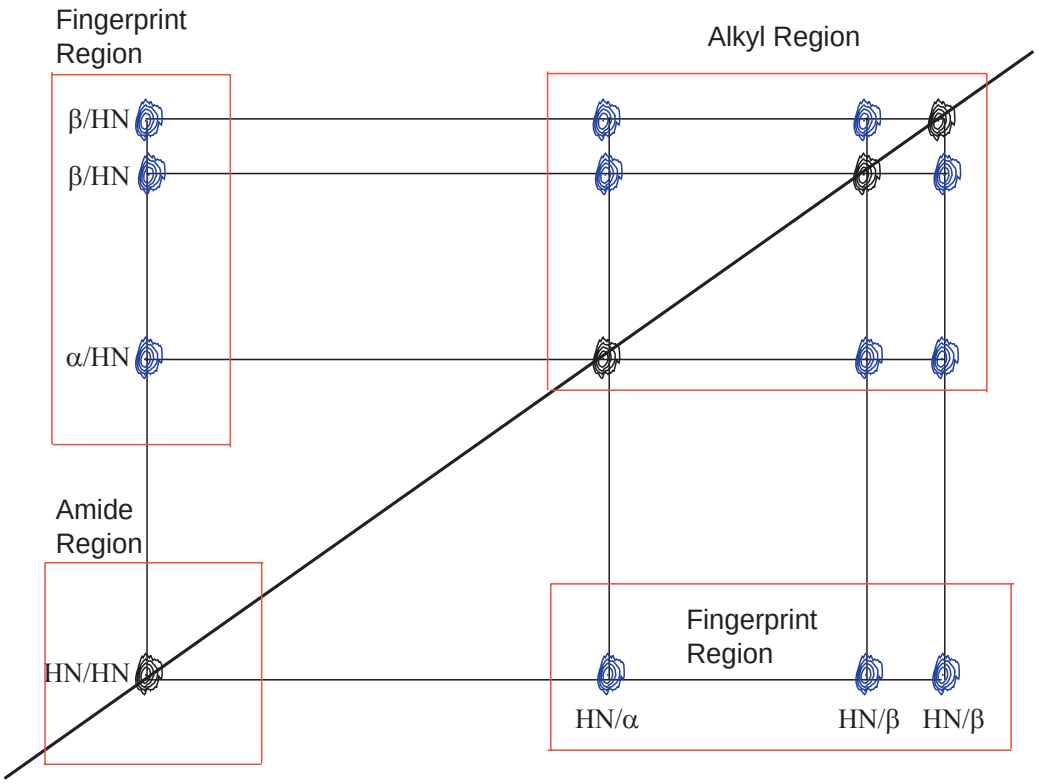


Figure 2.1.7.1: An illustration of how spin systems may be identified in a TOCSY spectrum. HN refers to the proton connected to the backbone amide nitrogen, α to the proton connected to the Cα in the side chain and β to the Cβ. Similarly, protons connected to carbons in longer side chains are labelled according to the carbon it is associated with. Thus, the peaks are labelled according to the protons that contribute to a particular peak in the same order as the magnetisation travels.

By tracing all the peaks aligned through a line vertically from an amide-shift on the left-hand side of the diagonal (the amide region), the potentially scalar coupled peaks are identified. To verify that the peaks are coupled and not the result of another spin system, the corresponding peaks originating from the α-proton and β-protons in the alkyl region are also traced. An array of peaks like the one in Figure 2.1.7.1 shows all possible connections in a TOCSY spectrum for an AMX spin-system.

Names of the different atoms are given in Figure 2.1.7.1, exemplified by the peaks in the fingerprint region. In short, a peak is named by stating the origin of the magnetisation first and the destination of

the magnetisation last. Thus, the peak identified as α/HN is a result of magnetisation travelling from the α-proton to the amide proton, whereas the HN/α peak results from magnetisation travelling in the opposite direction; from the amide-proton to the α-proton.

### 2.1.8 NOESY (Nuclear Overhauser Effect SpectroscopY); the Nuclear Overhauser Effect

Sequential assignment is completed by correlating protons that are close in space [86]. NOESY relies on the nuclear Overhauser effect (NOE) that links nuclei that interact through space, and are less than 5Å apart [77].

Nuclear Overhauser Effect

The intensity of an NMR-signal is proportional to the difference in population of two energy levels where nuclear resonance transition occurs [79]. Dipolar coupled nuclei do not relax independently [86]: Cross-relaxation (relaxation between two spins) leads to magnetisation transfer between the spins and the resulting change in intensity is known as the nuclear Overhauser effect [90]. The rate of cross-relaxation, and thus the NOE enhancement factor, is among others dependant on the distance between the coupled nuclei and the motion of the molecule [90, 95]. Cross relaxation leads to changes in the population of the two coupled spins. This change of population can again lead to peaks with stronger/smaller intensities as a result of the NOE-effect. Equalizing the population difference (saturation) of a spin S strongly affects the longitudinal magnetisation of a coupled spin I caused by the increased population difference of this spin [86, 90, 95]. The system is now not in a state of equilibrium and returns to equilibrium through relaxation mechanisms, mainly through dipole-dipole interactions [95]. NOE and dipole-dipole relaxation mechanisms are thus closely connected [79]. During cross relaxation, magnetism is transferred through dipolar couplings between the interacting nuclei, and the intensity of NOE-peaks is inversely proportional to the internuclear distance to the sixth power [77]. The intensity of the peaks is evaluated by integration, and classified as being strong, intermediate or weak [77].

In a 2D NOE experiment, such as NOESY, positive magnetisation is exchanged between nuclei [90]. By eliminating what peaks are due to scalar coupling (from comparison with the scalar coupled techniques described above), sequential assignment and distance constraints are derived from the remaining dipolar NOE-peaks [90]. Mainly three connectivities are utilized in this respect: HN-HN,

HN-Hα protons and HN-Hβ protons [90]. As a rule of thumb, the strongest NOE cross-peaks come from protons on sequentially neighbouring residues [80, 94].

For molecules with long correlation time τc, strong positive cross-peaks arise due to NOE-enhancement [90, 95]. Although curvacin A is a small polypeptide, it is associated with micelles (or membranes in biological environments). Curvacin A moves together with the micelles and hence behaves as if it was much larger. This results in slow rotational motion of the molecule and long correlation time [80, 90]. We therefore expect positive cross-peaks in the NOESY-spectrum of curvacin A.

Sequential assignment of the spin systems; NOESY

After having identified the spin systems, the next step is to sequentially assign the spin systems by identifying the intraresidual NOE couplings. This information is provided by a NOESY spectrum. Figure 2.1.8.1 schematically illustrates how sequential NOEs are located in a NOESY spectrum.

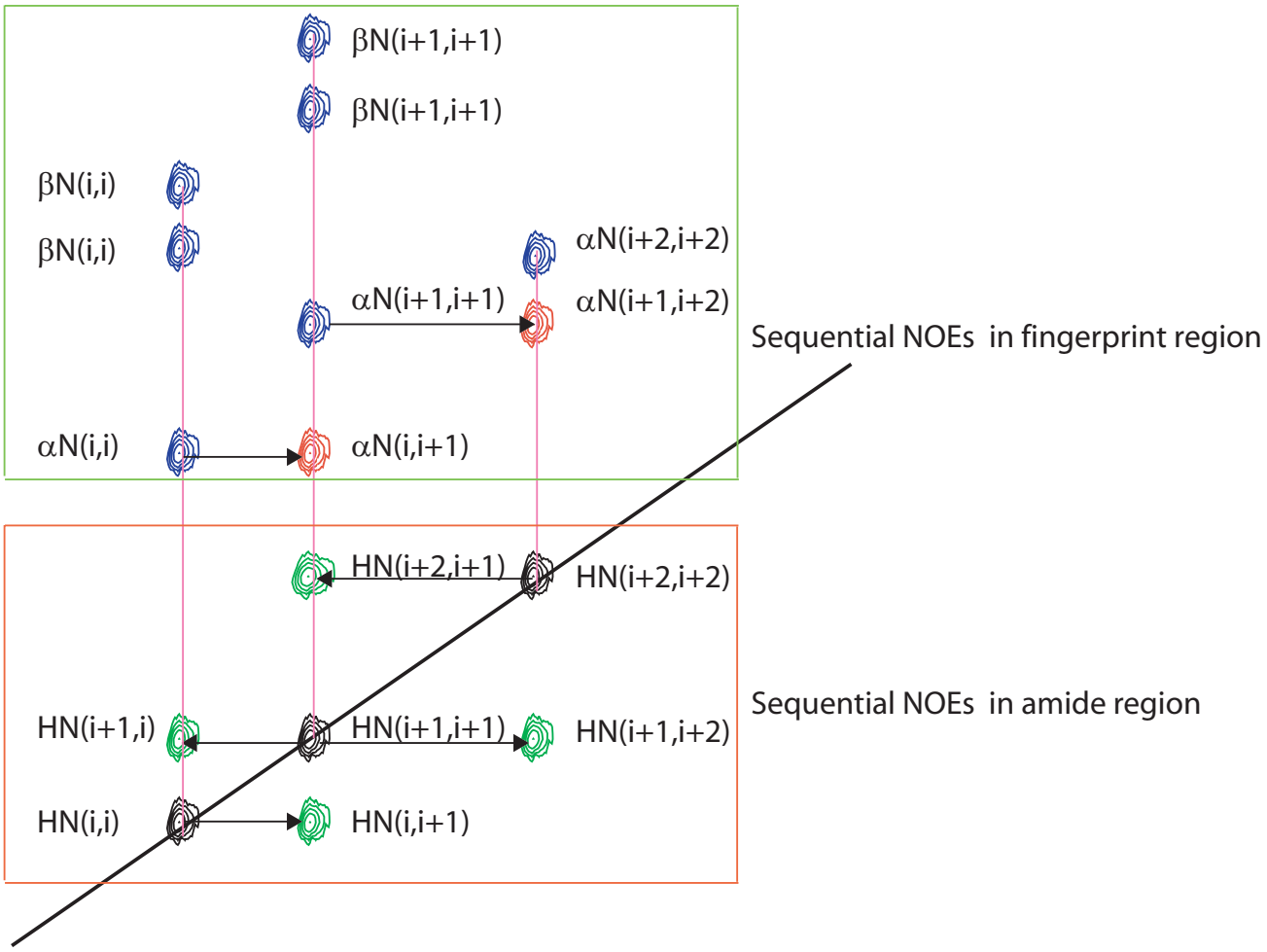


Figure 2.1.8.1: Illustration of how interresidue NOE connections in a NOESY spectrum can be identified. Blue and black peaks are intraresidue peaks already identified in the TOCSY spectrum. The red and green peaks are interresidue α-to-amide and amide-to-amide peaks, respectively. The horizontal arrows indicates how the cross peaks are read out. This is especially important in the fingerprint region, where the position of an interresidual NOE peak is sequence specific. See the text below for details.

In the fingerprint region, sequential couplings between side chain protons are identified. Starting at the α-proton of residue i and moving horizontally along the $\omega_2$-axis one should find a strong peak along the amide shift of residue i+1. This is the interresidual NOE peak between the α-proton of residue i and the amide proton of residue i+1. Following in the same manner, starting at the α-proton of residue i+1 and moving horizontally along the $\omega_2$-axis one should find a strong interresidual NOE peak along the amide shift of residue i+2.

In the same manner, cross peaks from β-protons to amide protons and other side chain protons supports the connection of the spin systems established by tracing cross peaks between α-protons and amide protons (not shown in Figure 2.1.8.1): Starting at the β/γ-proton of residue i and moving horizontally along the $\omega_2$-axis one should find a peak along the amide shift of residue i+1. This is the interresidual NOE peak between the β/γ-proton of residue i and the amide proton of residue i+1. Interresidual cross peaks between the amide proton of a residue and side chain protons other than the α-proton of a neighbouring residue are usually weaker than the cross peak between α-proton and amide proton on neighbouring residues.

One could also alternatively search for sequential connections in the opposite direction: Moving vertically up or down from the α-proton of residue i, one should find a strong interresidual NOE peak. By looking horizontally along the $\omega_2$-shift of this interresidual peak one should find the α-proton of residue i-1. This can be visualized by following the arrows from residue i+1 to i in Figure 2.1.8.1 in the opposite direction as to what is drawn.

Usually, a string of three consecutive spin systems is enough to determine the sequential location of these spin systems. However, sometimes more than three consecutive spin systems have to be identified to pick out a unique sequence. The aim is to locate a string of consecutive spin systems that can only be positioned one place in the sequence, and if there is more than one possible position for the string of spin systems identified (irrespectively of the number of consecutive spin systems), neighboring spin systems must be located. The number of residues in the molecule, type of spin systems represented in the sequence and their distribution along the sequence determines how many consecutive spin systems are needed to assign that particular sequence with confidence.

In the amide region, connections can be identified as described in Figure 2.1.8.1. Amide proton to amide proton on both sides of the diagonal may be identifiable. One cannot determine the direction of the connections from the backbone connections, only that the spin systems are connected sequentially. These connections are used methodically to verify the sequential connection found by tracing the cross peaks in the fingerprint region as described above. They are also important as distance constraints used for structure calculation in subsequent steps.

Aromates:

There are four aromatic residues in curvacin A; two tryptophans (W16 and W34) and two tyrosines (Y4 and Y9). The protons on the aromatic rings have characteristic topology, and correlate with the aliphatic side chain protons in NOESY spectra. By connecting the ring spin systems and the aliphatic AMX-spin systems identified in a TOCSY spectrum by the use of a NOESY spectrum, the aromatic residues can be singled out from the rest of the AMX-spin system residues.

### 2.1.9 Heteronuclear experiments:

In heteronuclear experiments, either $^{13}C$ or $^{15}N$ are used in combination with $^{1}H$ in 2D heteronuclear experiments in addition to the above homonuclear experiments. One thus obtains $^{1}H/^{13}C$ and $^{1}H/^{15}N$ 2D spectra.

In heteronuclear experiments of coherence transfer, rf pulses are applied at more than one frequency to manipulate two or more nuclei [86]. Utilizing more than one nucleus greatly improves the resolution in as much as overlapping proton resonances can be separated with respect to the different $^{13}C$ or $^{15}N$ chemical shifts associated with them, and finally aids the assignment procedure by simply contributing correlation between different species [90].

By adding another incremented dimension and another mixing period, the heteronuclear experiment can be expanded from 2D- to a 3D-experiment. (For example a $^{1}H/^{1}H/^{13}C$ 3D spectrum)

### 2.1.10 $^{13}C$ and $^{15}N$ HSQC (Hetronuclear Single Quantum Correlation)

HSQC makes use of a pulse sequence called INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) to transfer magnetisation from the sensitive proton nucleus to the insensitive $^{13}C$- or $^{15}N$-nucleus via direct $^{1}J$-couplings during the mixing period. The direct, or observed, F1 dimension is thus obtained from the insensitive nucleus whereas the indirect F2 dimension still stems from the proton. This creates a 2D heteronuclear spectrum. It is possible to run experiments where the direct dimension is read out on the insensitive nucleus ($^{13}C$-or $^{15}N$), but indirect (or proton) detection,

where the magnetisation is transferred back to the sensitive proton before it is read out, is used whenever possible in order to maximize sensitivity [86].

HSQC transfers correlation through H-C and H-N bonds. This is thus a $^{1}$J-coupled correlation technique, and can be used in sequential assignment of a molecule. For carbon, one can expect correlation peaks throughout the entire carbon-bonded side chain of an amino acid. For nitrogen, mainly the amide backbone cross peaks are of interest, but for some of those amino acids with amides in their side chains, additional peaks are expected.

It is possible to perform structural determination from homonuclear experiments [96], as long as the proteins are of modest size (molecular weight less than 10 000). Even so, it is useful to run J-correlated experiments such as HSQC on natural abundance of $^{13}$C and/or $^{15}$N to exploit the additional chemical shift-data from such experiments [97].

### 2.1.11 Structure calculation, the CYANA-program:

When having identified (preferably) all spin systems and labelled them according to the amino acid sequence using scalar coupling, additional information yielding distance constraints such as NOE-cross peaks, volumes of these peaks, three-bond J-couplings and chemical shifts is used to calculate a plausible 3D-structure [86]. The success of such a strategy for structure calculation depends on an as complete list of chemical shift assignments for protons as possible [98]. It is doable for an experienced spectroscopist to calculate structures manually for small molecules, but it is cumbersome and time-consuming. Rather, one of several computerized model-simulating programmes is used. Here, CYANA (Combined assignment and dYnamics Algorithm for NMR Applications) is used for automated NMR structure calculation [99, 100].

Basically, what CYANA does is through several refining cycles of conformational computations to reach a structural model with the least deviation from the experimental data [101]. The calculated structure with minimum energy is the structure that has maximum compatibility with the experimental data. This iterative process utilizes preliminary structures from earlier cycles to reduce ambiguities of cross-peak assignments. Additional selection criteria for cross-peaks and NOE-

assignments are added after each cycle, yielding better structures in the process. Cycle one is thus the most crucial one, being the only for which a previous structure is not available [99, 101].

The key conformational data-input are upper distance limits derived from nuclear Overhauser effects (NOEs) [100].

A difficulty with NMR is the unreliability with which all information can be accurately assigned and labelled. As mentioned above, distances constraints from NOEs are important for structure calculation. However, a majority of these NOE cross-peaks cannot be unambiguously assigned, i.e. there is more than one possible assignment for most of the peaks. CYANA overcomes this obstacle in several ways. First, all ambiguous cross-peaks are treated as the superposition of the signals from each of its multiple possible assignments, and weighing them inversely proportionally to the corresponding interatomic distance. As long as all the correct assignments are among the assignments that are selected, the distance constraint is never falsified, even though incorrect assignments are also included [100].

Constraint combination takes the above a step further; assignment from different, often unrelated, cross-peaks are combined to generate distance restraints.

Furthermore, the concept of network anchoring is a powerful tool for eliminating false conclusions. The basis of network anchoring are that each NOE assignment must be connected to all other assignments in a network without contradictions. This ensures that “lonely” constraints from e.g. unstructured coil-regions are not falsely forced into a defined structure [100].

The actual calculation performed by CYANA, is carried out by defining a target function that measures the agreement between a structure and a given set of constraints [99, 101]. A low target function value reflects a conformation that satisfies the constraints more closely than another with higher function value. Thus, best agreement between constraints and a calculated 3D structure is defined as the set giving the lowest target function value.

Simulations done with different chemical shift assignments left out show that at least 90% of the shifts must be assigned for a satisfactory result. If some “essential” assignments that contribute

strongly to the structure are left out, a higher percentage might be necessary for correct structure calculation [98, 100].

## 2.2 Other methods used in this thesis:

### 2.2.1 Circular Dichroism, CD:

Biopolymers such as proteins rotate the plane of plane polarized light, and a CD experiment measures the difference in absorption of left and right circularly polarized light [102]. The difference, or ellipticity, $\Delta A$ between left and right circularly polarized light is plotted as a function of wavelength in a CD spectra. Secondary structure elements such as random coil, $\alpha$-helix and $\beta$-sheet affect the spectra and give characteristic curves [102]. The difference in the curves are mainly due to asymmetry in the polypeptide backbone as a result of different conformations [102]. A CD-spectrum can be used to quantitatively estimate the relative percentage of the different secondary structure elements in a sample, but cannot be used to determine where in the sequence these elements are [102].

Even if not giving the whereabouts of structural elements, such a spectrum can give valuable information as to whether a polypeptide is structured at a given set of conditions and at what concentrations of these conditions optimal structuring is observed.

### 2.2.2 Matrix-Assisted Laser Desorption Ionisation Time of Flight Mass Spectrometry (MALDI-TOF MS):

Two major applications of mass spectroscopy make it an attractive technique for analysis of bioorganic molecules: the ability to elucidate primary structure, and mass determination of high accuracy [103]. In addition, the method displays very high sensitivity. In the work presented here, the latter application has been used to verify the purity and mass of curvacin A in fractions from reverse-phase chromatography, and only this aspect of MS will be considered here.
The Laser Desorption Ionisation -technique (LDI) is well suited for analysis of large proteins and peptides varying from several hundred to several thousand Daltons [103], and differentiates species based on their different mass:charge (m/z) ratio in an MS spectrum [104]. When in an applied magnetic field, charged species with mass m and velocity v will be deflected by a degree dependant on mass [102]. When the signal is plotted as a function of time, peaks corresponding to ions of

different mass will appear at different times [104]. Thus, accurate mass determination can be achieved by the known relationship between TOF and m/z-ratios [103, 105].

Briefly described, in MALDI TOF a sample is ionised by UV-irradiation to charged particles and accelerated by an electric or magnetic field. The ions are all accelerated to the same velocity and hence only the mass has an impact on deflection. The time an ion require to arrive at the detector is measured (time of flight), and m/z is calculated from this [104, 105]. Masses can be determined with accuracies up to ±0,01% (i.e. ±0,5 Da for curvacin A) [102, 104].

# 3 Materials and Methods:

## 3.1 Production of curvacin A

Two procedures for producing curvacin A in amounts sufficient for NMR-analysis were investigated: 1) Cloning of the curvacin A-encoding gene into an *Eschericia coli* (*E. coli*) over-expression system (described in section 3.1.1 below) and 2) Production of curvacin A from its natural producer, *Lactobacillus curvatus* LTH1174 (see section 3.1.2 below).

### 3.1.1 Cloning and over-expression of curvacin A in *E. coli*-cells

Briefly, the curvacin A gene isolated from LAB LTH1174 was ligated into a pET 3a plasmid (Novagene) and transformed into *E. coli* for over-expression (Figure 3.1.1.1). The plasmid (pET 3a-vector + curvacin A gene) was first transformed into *E. coli* DH5α–cells for cloning of the curvacin-encoding plasmid. Subsequently, the plasmid was transformed into *E. coli* BL21(DE3)pLysS–cells (Stratagene) for protein expression.

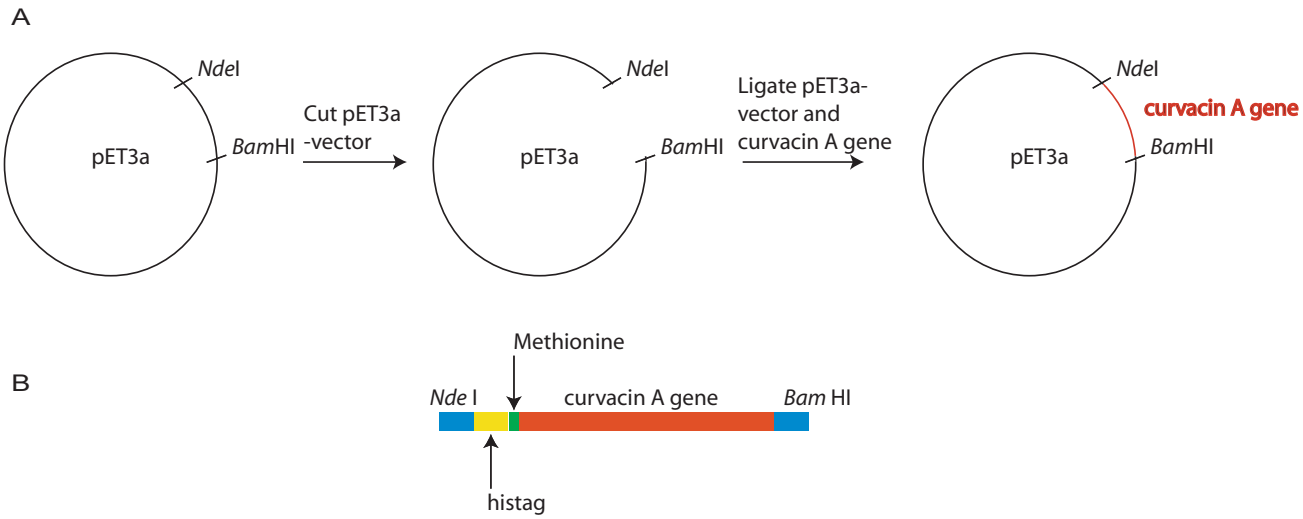


Figure 3.1.1.1: Description of constructed plasmid for over-expression of curvacin A in *E. coli.* **A** Schematic drawing illustrating the construction of pET3a plasmid with incorporated curvacin A gene. The curvacin A gene was incorporated by cutting the pET3a plasmid with restriction enzymes *Nde* I and *Bam* HI and subsequently ligating in the curvacin A gene with complementary ligation sites. **B** Detailed drawing of the inserted curvacin A gene (red) and the primers attached. The primers are the components flanking the curvacin A gene on either side. Blue: Cutting sites for the restriction enzymes *Bam*HI and *Nde*I. Yellow: A histidine tag (histag). Green: C-terminal methionine residue of curvacin A-forward primer.

Table 3.1.1.1 shows the primers used in the PCR-reaction.

Table 3.1.1.1: The two oligonucleotides used in the PCR-reaction for amplification of plasmids isolated from *Lactobacillus curvatus* LTH 1174 (Curvacin A forward and Curvacin A reverse), and the two oligonucleotides used in the mutagenesis of methionine 30 to alanine (Curvacin A M30R and Curvacin A M30F).

| **Name** | **Oligonucleotide sequence** | **Notes** |
|---|---|---|
| Curvacin A forward | 5'-CGT-CTC-TAC-ATA-TGT-CGC-ATC-ATC-ATC-ATC-ATC-AC**A-TG***G-CAT-GAT-CTT-ATG-GCA-ACG-GTG-TTT-AC*-3' | Histag underlined, methionine codon in bold and underlined, and beginning of curvacin A gene in italic |
| Curvacin A reverse | 5'-CTA-GAC-GGA-TCC-TTA-TTA-**TGC**-TCC-AGC-TAA-ACC-ACT-A-3' | BamHI site underlined and M41A mutation codon in bold and underlined. |
| Curvacin A M30R | 5'-CCA-GCC-GCT-AAT-**TGC**-ACC-ACC-AAT-AAT-ACT-TTG-3' | M30A mutation codon in bold and underlined. |
| Curvacin A M30F | 5'-CCA-AGT-ATT-ATT-GGT-GGT-**GCA**-ATT-AGC-GGC-TGG-3' | M30A mutation codon in bold and underlined. |

Cleavage of the fusion peptide and liberation of mature curvacin A were done by cyanogen bromide (CNBr) essentially as described before [106]. CNBr specifically cleaves at the C-terminal side of methionine residues. Since curvacin A contains two methionine residues, it was necessary to remove these by site directed mutagenesis in order to get full-length curvacin A-peptide after CNBr-cleavage. Thus, the curvacin A reverse- oligonucleotid was designed to change the C-terminal methionine (residue 41) of curvacin A to alanine, and another set of oligonucleotides were constructed to susbstitute methionin 30 to alanine. The methionine residue needed for CNBr-cleavage was encoded by the curvacin A forward-oligonucleotide which also encodes for the N-terminal histag (Figure 3.1.1.1B). This methionine will, before purification of curvacin A on the Äkta FPLC system, provide the cutting seat for cyanogen bromide (CNBr) cleavage of the histag to release the mature protein.

Cloning of curvacin A into an *E. coli* BL21(DE3)pLysS-vector for over-expression:
*Lactobacillus curvatus* LTH 1174 [48] was grown overnight in 10 ml of MRS. After sentrifugation, the cell pellet was resuspended in QiAgene cell resuspension solution and lysozym treated before cell lysis and completion of the QiAgene procedure for isolation of plasmids mini prep, as described in the producer's User Manual (QiAgene). The isolated plasmids were analysed on a 1% agarose gel and the curvacin A gene was then amplified by using the oligonucleotid set shown in Table 3.1.1.1 and a standard PCR-reaction on a MJ Research PTC-200 Peltier Thermal Cycler.

The PCR-product and the pET3a plasmid were cut with restriction enzymes *Nde*I and *Bam*HI (Fermentas), and purified on a 1% agarose gel and by using the GFX$^{TM}$ -kit as described in the GFX$^{TM}$ PCR DNA and Gel Band Purification Kit Manual from Amersham Biosciences.

pET 3a was furthermore treated with the enzyme Calf Intestine Alkaline Phosphatase (CIAP), also from Fermentas, to minimize the amount of vector that re-ligated with itself. CIAP specifically cleaves off the phosphate group at both 5' and 3' end of DNA/RNA- backbones. Hence, ligation of a 5'or 3' end of pET3a together with the complementary strand from curvacin A with intact phosphategroup is more likely than re-ligation of the vector itself.

The purified curvacin A PCR-products were then ligated into the pET 3a-vector (Novagene) (Figure 3.1.1A), using the procedure described in the Fermentas Manual.

The resulting ligmix was thereafter transformed into chemo-competent *E. coli* DH5α-cells and grown over night on LB agar plates containing 100 μg ampicilin/ml (Biolabs Manual). Colonies were picked from agar plates the following day, and left to grow overnight in LB-medium, supplemented with ampicillin. Plasmids with incorporated curvacin A gene were identified on a 1% agarose gel and by using PCR and the oligonucleotides shown in Table 3.1.1.1.

In order to remove the last methionine residue in curvacin A, the ATG codon encoding it was substituted to GCA encoding alanine by using the Quick Change Site-Directed Mutagenesis kit (Stratagen) and the oglionucleotides shown in Table 3.1.1.1.

The mutated plasmids were then transformed into *E. coli* DH5α-cells with selection on agar plates containing 100 μg ampicillin/ml. After harvesting the colonies, the plasmids were isolated (QiA gene Miniprep Manual) and the DNA was sequenced using MegaBase DNA analysis system. Plasmid with the desired insert was transformed into *E. coli* BL21(DE3)pLysS-cells and grown on agar plates containing 100 μg ampicillin/ml and 30 μg chloramphenicol/ml for selection.

Induction and over-expression of curvacin A in *E. coli* BL21(DE3)pLysS:

Cultures grown over night in 3 ml of LB-medium containing chloramphenicol (30 μg/ml) and ampicillin (100 μg/ml) were centrifuged. The cells were washed twice with fresh LB-medium to remove the enzyme penicillinase that inactivates ampicillin. The cells were transferred to 100 ml of LB-medium containing ampicillin and chloramphenicol. The cultures were left to grow to optimal density (OD 0.5) before isopropyl-beta-D-thiogalactopyranoside (IPTG) (Fermentas) was added. The *E. coli* BL21(DE3)pLysS expression system contains fragments of the bacteriofag T7 DNA that among other contain the T7 RNA polymerase. IPTG indirectly induces the T7 RNA polymerase, which then transcribes the target gene (curvacin A). After inoculation for about two hours, the cells were lysated by three alternative chemicals; BugbusterTM/Benzonase, 6 M Guanidine HydroChloride (Gu-HCl) (Buffer A, Table 3.1.2) and 8 M urea (Buffer B, Table 3.1.2). The culture was centrifuged and the cleared lysate purified on a Histag binding column, Ni-NTA His-Bind® Resin (Novagene) as described in the producer's manual.

Table 3.1.2: Buffers used upon purification of curvacin A from *E. coli* BL21(DE3)pLysS-cells

| **Buffers** | **Content** |
|---|---|
| A | 6 M Gu-HCl; 0.1 M $NaH_2PO_4$; 0.01 M TrisCl; pH=8 |
| B | 8 M urea; 0.1 M $NaH_2PO_4$; 0.01 M TrisCl; pH=8 |
| C | As B, but pH=6.3 |
| D | As B, but pH=5.9 |
| E | As B, but pH=4.5 |

The lysates were mixed with a 50% Ni-NTA His-Bind slurry (Novagene) before application on the column.

The column was washed with 2x4 ml buffer C (Table 3.1.2) before the recombinant protein was eluted with 4x0.5 ml of buffer D and 4x0.5 ml of buffer E (Table 3.1.2), respectively. The combined fractions of protein eluated with buffers D and E was then ready for CNBr-cleavage and purification on the Äkta FPLC-system.

### 3.1.2 Production of curvacin A by its natural producer

For production of curvacin A by its natural producer, *Lactobacillus curvatus* LTH1174, the producer was seeded from a -80$^{o}$C stock to 10 ml MRS broth (Oxoid) and cultured over night at 30 $^{o}$C. This 10 ml culture was then transferred to 2 L of MRS broth which was cultured over night, and a stationary phase culture was obtained, from which curvacin A was isolated as described in section 3.3 below.

## 3.2 Bacterial Strains and Bacteriocin Assay

By using a mircotiter plate assay system, bacteriocin activity was measured as described previously [59, 107, 108]. Each well of the microtiter plate contained bacteriocin fractions at twofold dilutions and the indicator strain *Lactobacillus sake* NCDO 2714 in 200 µl of MRS broth. The microtiter plate cultures were incubated over night at 30$^{o}$C and by thereafter measuring spectrophotometrically at

600 nm with a Tecan Sunrise microplate reader, the growth inhibition of the indicator organism was determined.

### 3.3 Isolation of curvacin A by ion exchange and reverse phase chromatography

Isolation of curvacin a was carried out by using a newly established purification protocol [109].

Overnight cultures of curvacin A were applied directly on a 3 ml Sepharose Fast Flow cationic exchange column. The column had been equilibrated with 100 ml of 20 mM sodium phosphate, pH 5.2. About 0.5 l of over-night culture was applied per run. After application, the column was first washed with 100 ml of 20 mM sodium phosphate, pH 5.2, and then washed with 30 ml of 0.15M NaCl in 20 mM sodium phosphate, pH 5.2. The sample was eluted with 40 ml 1 M NaCl in 20 mM sodium phosphate, pH 5.2.

The eluent obtained from the cationic exchange column was applied to a 3 ml RESOURCE RPC reverse phase column on the Äkta purifier-system. Curvacin A was then eluted from the column with a linear elution gradient from 0% (v/v) 2-propanol, 0.1% (v/v) TFA to 50% (v/v) 2-propanol, 0.1% (v/v) TFA. The elution volume was 60 ml, and the flow speed 1.5 ml/min. The absorbance at 280 nm as a function of ml eluent was recorded.

### 3.4 Analysis of curvacin A on a reverse phase column using the SMART-System (Micro-Preparative/Analytic Chromatographic System)

To verify the degree of purity of isolated curvacin A, the SMART system with an analytical (μRPC $C_2/C_{18}$ SC.2.1/110) reverse phase column was used. A linear gradient from 0% (v/v) 2-propanol, 0.1% (v/v) TFA to 30% (v/v) 2-propanol, 0.1% (v/v) was applied. The eluent volume was 2 ml, and the flow speed 100 μl/min. The absorbance at 214, 254 and 280 nm as a function of ml eluent was recorded.

### 3.5 Analysis of purified curvacin A by MALDI-TOF MS

To verify the identity and purity of the isolated curvacin A, MALDI-TOF MS with a reflector mode was used. The matrix consisted of a 1:1 ratio of 3% TFA and acetonitril, saturated with a-cyano-4-hydroxycinnamic acid.

### 3.6 CD Measurements of curvacin A

CD spectra were recorded by using a Jasco J-810 spectropolarimeter (Jasco International Co., Ltd., Tokyo Japan) calibrated with ammonium D-camphor-10-sulfonate (Icatayama Chemicals, Tokyo Japan). All the measurements were performed by titrating curvacin A at a concentration of 0.15 mg/ml with increasing concentrations of DPC (final DPC concentration of 16 mM; CDN isotopes, Quebec Canada). Measurements were performed at various temperatures, from 20-55 $^{o}$C, using a quartz cuvette (Starna, Essex England) with a path length of 0.1 cm. Samples were scanned 5 times at 50 nm/min with a band width of 1 nm and a response time of 1 s, over the wavelength range 190 to 260 nm. The α-helical content of the various peptides under different solvent conditions was determined by a single point method using the mean residual ellipticity at 222 nm ($[\theta]_{222}$) [110, 111] and the equation below:

$$f_H = [\theta]_{222} / (-40{,}000(1-2-5/n))$$

where $f_H$ is the α-helical content and n is the number of peptide bonds.

### 3.7 NMR Sample Preparation for curvacin A

For NMR structure elucidation, 6.8 mg of curvacin A were dissolved in 750 μl of 350 mM deuturated DPC (CDN Isotopes), 10% $D_2O$ (Cambridge Isotope Laboratories) and 0.1% trifluoroacetic acid. The final concentration of curvacin A in the sample was 2.1 mM.

### 3.8 NMR Spectroscopy of curvacin A

The NMR spectra of curvacine A were obtained at 25$^{o}$C, 30$^{o}$C and 35$^{o}$C on an 800 MHz Varian INOVA 800 NMR spectrometer with 4 channels and 5 mm 1H {13C, 15N} pfg probe. TOCSY [112] and NOESY [113, 114] experiments included in the biopack were performed to assign the molecule. Spectra with mixing times of 48 and 64 ms were acquired for the TOCSY pulsesequence and 100 and 150 ms were acquired for the NOESY pulsesequence. Watergate water decoupling was applied in the TOCSY and NOESY experiments [115]. 1024K complex data points where obtained in the direct dimension and 512 in the indirect dimension. A sinbel function was applied to data and it was zero-filled to the double size prior to Fourier transformation. All post processing was obtained with the application of NMR Pipe [116], while spectral assignments and integration were done with the application of SPARKY (T. D. Goddard. and. D. G.Kneller, University of California, San Francisco). NOE restrictions were obtained from the structure assignment of the molecule.

### 3.9 Restraints and structure calculation

Dihedral angle restraints were obtained by using the TALOS program [117] on the chemical shifts values. The NOESY spectrum of the molecule was manually assigned using standard methods [114]. A total of 567 NOE distance constrains were obtained after removal of those that can not be violated or where found several times, out of which 244 were interresidue and 323 were intraresidue NOE restraints. 24 long-range NOE restrains distances larger than 5 residues were observed. The structure was calculated and annealed by the application of the structure calculation program CYANA and the anneal function. After several rounds of calculations, 6 hydrogen bonds (E21 to S25, I31 to A35, S32 to S36, G33 to G37, W34 to L38 and A35 to A39) were introduced in the final annealing of the structure.

A total of 100 structures were calculated out of which the 20 best structures were selected for further evaluation. The target function gave an average of $1.71 \cdot 10^{-2} \pm 9.43 \cdot 10^{-3}$. The structures were visualized with the MOLMOL program [118].

# 4 Results and Discussion:

## 4.1 Growth and purification of curvacin A:

### 4.1.1 Purification of curvacin A from complex culture media

Curvacin A was purified by chromatography on a SP-Sepharose Fast Flow cation exchange column followed by chromatography on a RESOURCE RPC column (Figure 4.1.1A), all according to the purification method developed by Uteng et al. [109].

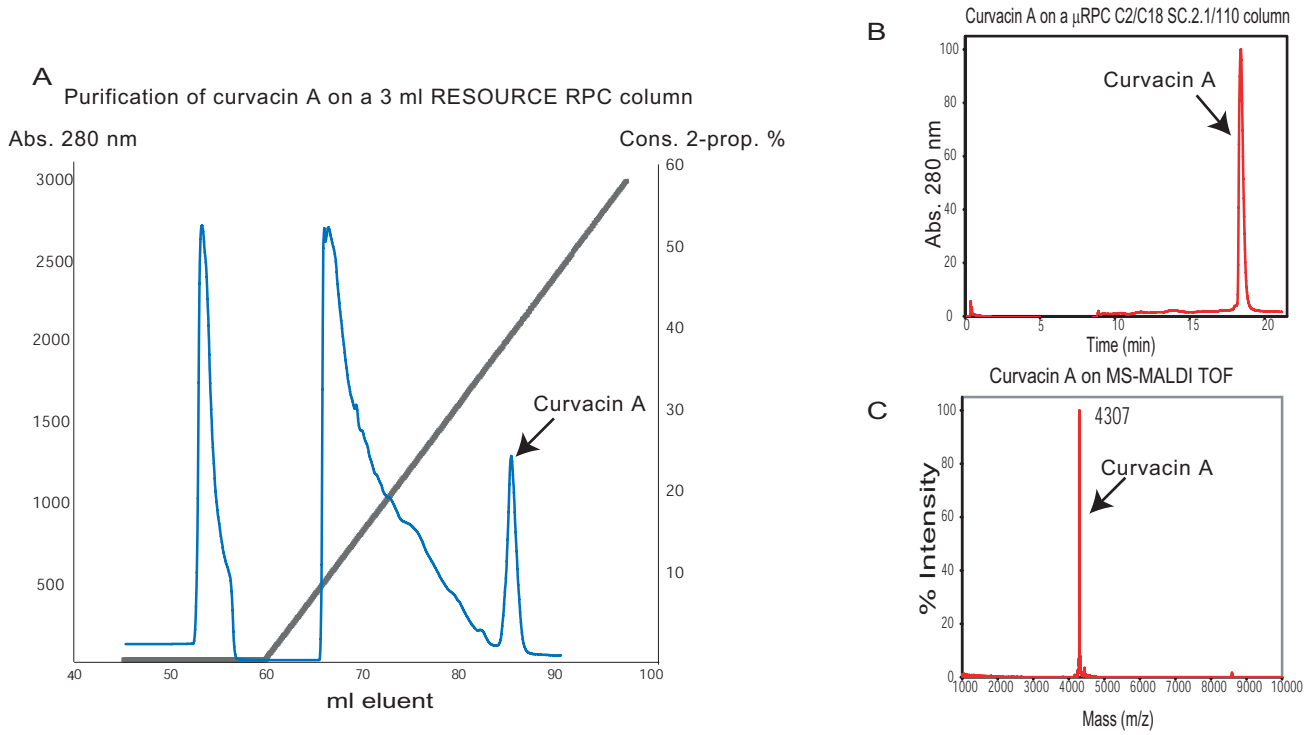


Figure 4.1.1: **A** Chromatogram resulting from purification of curvacin A on a 3 ml RESOURCE RPC column using the Äkta-system. The blue line is absorbance recorded at 280 nm (left vertical axis) whereas the black line shows the linear gradient of 2-propanol (v/v), 0.1% (v/v) TFA used to eluate curvacin A (right axis). Both absorbance and concentration were recorded as a function of ml eluent. **B** Chromatogram resulting from analytical analysis of purified curvacin A on a µRPC $C_2/C_{18}$ SC.2.2/110 column using the SMART-system. **C** Mass spectrum of purified curvacin A obtained by MS-MALDI TOF.

Chromatogram obtained by purification on the RESOURCE PRC reverse phase column using an Äkta purifier (Amersham Pharmacia Biotech) is presented in Figure 4.1.1A. The bacteriocin appeared as a major optical density peak upon reverse phase chromatography. The chromatogram in Figure 4.1.1A shows that curvacin A eluted when the concentration of 2-propanol reached 42%.

The purity of the fractions of curvacin A was subsequently analysed by reverse phase chromatography as described in section 3.4 and by mass spectroscopy as described in section 3.5. The chromatogram obtained upon chromatography using an analytical (μRPC $C_2/C_{18}$ SC.2.2/110) reverse phase column, is shown in Figure 4.1.1B. Figure 1.4.1B verifies the purity of the fractions obtained from reverse phase chromatography shown in Figure 4.1.1A as curvacin A eluated as the only major optical density peak (Figure 4.1.1B).

To ensure that the molecule in the purified sample was curvacin A and to verify the purity of the sample, the molecular mass of the peptide in the sample was determined by mass spectroscopy using a MALDI TOF-MS (Figure 1.4.1C). The molecular mass determined was in agreement with the theoretical molecular mass of curvacin A as calculated from the amino acid sequence. The mass spectrum also revealed that the two cysteines were covalently bonded in a disulfide bridge (the observed molecular mass is 2 Da lower than what would be expected if the cysteines were not bonded). The fact that the isolated peptide was active and that there is a great loss of activity under reducing conditions (where the disulfide bridges are no longer present in the molecule) [14] supports this conclusion. From this we concluded that the sample was pure, and that the purified molecule was curvacin A.

The overall yield of curvacin A from the overnight cultures was determined by measuring the absorbtion spectrophotometrically at 280 nm and using the equation

$$A = \varepsilon c l,$$

where A is the absorbtion measured, ε is the molar extinction coefficient (13940 $M^{-1}cm^{-1}$ for curvacin A), c is the concentration of the sample (the unknown parameter in our experiments) and l is the travelling path of the light (1 cm in these experiments). The purified fractions of curvacin A

from a known volume of starter culture were measured. From this, the yield was calculated to be on average 0.5 mg curvacin A per litre of starter culture.

A drawback of using the method described above to produce curvacin A is that production of labelled curvacin A is not practical. Molecules labelled with $^{13}C$ and/or $^{15}N$ may markedly simplify the NMR-spectra analysis (such as resonance assignment and subsequent structure calculation) [96]**.** Additional information from chemical shift, spin couplings and relaxation behaviour may furthermore be obtained by heteronuclear experiments [96]. The development of a procedure for producing labelled curvacin A was consequently attempted, and is described in the following section.

### 4.1.2 Over-expression of curvacin A in *E. coil*-cells

The curvacin A gene was introduced into an over-expression system as described in section 3.1 [48, 119]. The mutations of the methionines and the transformation into *E. coli* DH5α-cells were successful (data not shown). During the subsequent transformation into *E. coli* BL21(DE3)pLysS-cells for protein expression, we encountered unforeseen problems. Several trials of transformation were performed, where point-mutations from cysteine-encoding codons to stop-encoding codons, and loss (or partial loss) of the incorporated histag were among some of the problems encountered (data not shown). Due to these unforeseen problems, this track was abandoned. The procedure has successfully been reported for similar bacteriocins [37], and the source of our problems have not been identified. Instead, curvacin A was purified as described in 4.1.1 above. This was by-and-large done to ensure that a NMR- sample could be obtained within the allotted time.

Recently, the fusion gene of the lactococcin G α/β gene and the immunoglobuline-binding domain gene of streptococcal protein G (GB1 domain) was introduced into *E. coli* (Per Rogne, personal communication), by using a recently developed expression system [120, 121]. The yield of fusion protein was high (Per Rogne, personal communication). The fusion protein was then cleaved with CNBr, yielding a mixture of the mature peptide and the GB1-domain. The mature peptide and the GB1-domain were separated by reverse phase chromatography. The purified labelled bacteriocin has since been analyzed by NMR (Per Rogne, unpublished results).

For future production of labelled curvacin A and/or other pediocin-like bacteriocins for NMR-studies, the development of this expression system should be pursued. Due to the methionine residues, an alternative cleavage procedure must however be applied for curvacin A or alternatively, the methionine residues in curvacin A may be mutated (see section 3.1 for details) [122].

## 4.2 Analysis of curvacin A using CD-spectroscopy

As mentioned in the introduction, pediocin-like bacteriocins are unstructured in water, but become structured upon interaction with membrane-mimicking entities such as DPC. To determine at what conditions we have optimal structuring of curvacin A (peptide:micelle-ratio and temperature), curvacin A-structuring was investigated by the use of CD-spectroscopy at various temperatures and concentrations of DPC-micelles (Figure 4.2.1).

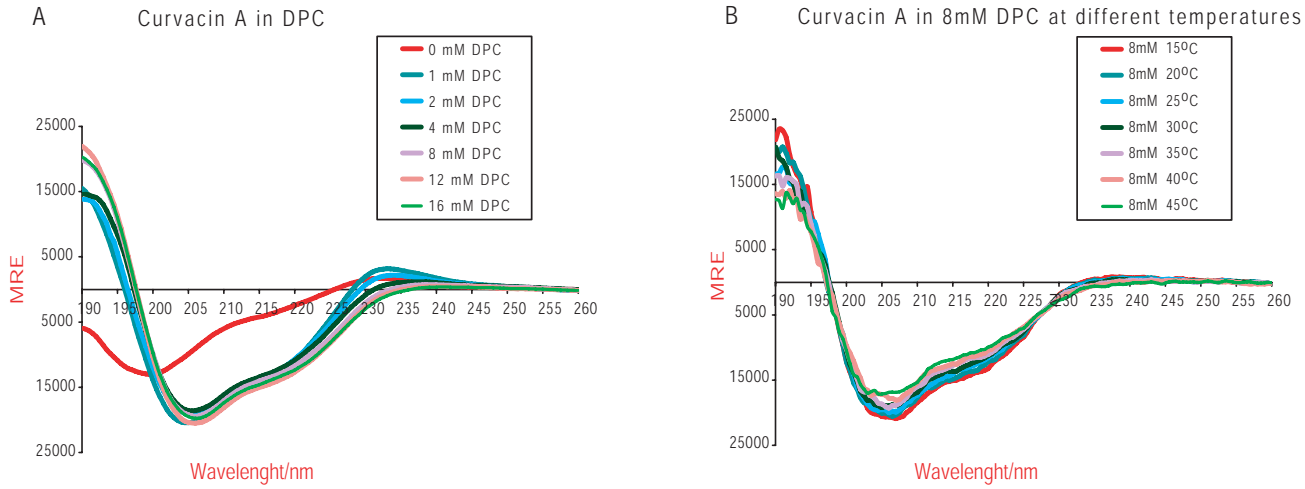


Figure 4.2.1 CD-spectra of curvacin A. Mean residual ellepticity (MRE) is plotted as a function of wavelength measured in nm. The concentration of curvacin A was fixed at 0.15 mg/ml. **A** Curvacin A dissolved in DPC at concentrations varying from 0 mM to 16 mM at 20°C. **B** Curvacin A dissolved in 8 mM DPC analysed for structuring at different temperatures.

Characteristic CD-spectra for α-helices have negative peaks at about 208 and 222 nm and a positive peak at about 192 nm [102, 123]. Figure 4.2.1A reveals that curvacin A attains α-helical-structuring upon interacting with DPC-micelles. Curvacin A is unstructured at 0 mM DPC, but becomes structured at DPC concentrations above the critical micelle concentration of 1 mM (the minimum

concentration of DPC in a solution if micelles are to be formed). Already at a DPC concentration of 2 mM, curvacin A displayed a CD-spectrum with characteristics typically found for α-helices. At about 8 mM DPC, the increase in the helical content of curvacin A comes to a rest within the uncertainty of the experiment (Table 4.2.1). Evidently, maximum structuring is achieved at 8 mM DPC, at the given concentration of curvacin A (0.15 mg/ml, equivalent to $3.5*10^{-2}$ mM), with 27-29 % helicity [110] (See section 2.2.1 for equation that determines % helicity from CD-spectra). This corresponds to 11-12 amino acid residues in helical conformation.

There is equilibrium between free peptide (P) and micelles (M) on one hand, and peptide bound to micelles (PM) on the other hand:

$P + M \leftrightarrow PM.$

The high concentration of micelles drives the equilibrium towards the bound state of peptide and micelle (PM), and hence structured peptide. It appeared from the CD-studies that at 8 mM DPC, curvacin A (at a concentration of 0.15 mg/ml) had close to optimal structuring (Table 4.2.1A). That is, most of the peptide is structured and associated with micelles. Higher micelle-concentrations would drive the equilibrium further towards the bound state. However, the marginal changes in structuring of curvacin A at higher concentrations of DPC would probably not produce changes in the NMR-signal (the % of structured peptide is at 8 mM already close to 100). Thus a concentration of 8 mM DPC was chosen for further CD-experiments.

In Figure 4.2.1B, 0.15 mg/ml curvacin A dissolved in 8 mM DPC has been titrated at temperatures from 15 $^{o}$C to 45 $^{o}$C. At all temperatures, the CD-spectra were typical for α-helical peptides. At higher temperatures (40 $^{o}$C and 45$^{o}$C), however, a loss of structuring was observed. This was especially pronounced in the area around 200-210 nm, where the intensity of the negative peak decreased. This indicates loss of secondary structure, possibly due to thermal instability of the three-dimensional structure. In the temperature range between 25 $^{o}$C and 35$^{o}$C, curvacin A has a stabile α-helical content, and the curves are equal within the uncertainty of the experiments.

Table 4.2.1 Statistical data from CD-spectroscopy of curvacin A. **A** Correlation between micelle concentrations and secondary structure of curvacin A [110, 111]. The peptide:micelle-ratio for the corresponding concentration of DPC is shown. Experiments performed at 20°C **B** Correlation between increased temperature in a sample with constant concentration of DPC = 8 mM and secondary structure of curvacin A [110, 111]. The concentration of curvacin A is fixed at 0.15 mg/ml throughout the experiments.

**A**

| DPC in mM at 20°C | Peptide:micelle-ratio at different DPC concentrations | % α-helical content of curvacin A |
|---|---|---|
| | | |
| 0 | ∞ | 4.5 |
| 1 | 2 | 23 |
| 2 | 1 | 23.5 |
| 4 | 0.5 | 24.5 |
| 8 | 0.25 | 27 |
| 12 | 0.15 | 29 |
| 16 | 0.1 | 29 |

**B**

| Temperature in °C of curvacin A in 8 mM DPC | % α-helical content of curvacin A |
|---|---|
| 20 | 27 |
| 25 | 30 |
| 30 | 29 |
| 35 | 27.5 |
| 40 | 26 |
| 45 | 25 |
| 50 | 24 |
| 55 | 23 |

Under the conditions of the CD-experiments above, DPC forms spherical micelles [124, 125]. Molecular modelling has shown that an average of 54 molecules of DPC form one micelle [124]. The relative amounts of peptide and DPC in the CD-experiments, shows that a peptide:micelle-ratio $\leq 0.25$ is needed for optimal structuring when the DPC is $\geq$8 mM. The relative large amount of DPC compared to curvacin A needed to obtain optimal structuring suggests that a molecule of curvacin A is not permanently associated with a particular micelle, but rather fluctuates between a micelle-associated structured state and an unstructured state.

## 4.3 NMR structure calculation of curvacin A

For a NMR-sample, several mg of curvacin A are necessary to obtain good spectra. Here, 6.8 mg of curvacin A in a total sample volume of 140 µl was used. From the CD-experiments we found that the DPC concentration had to be ≥8 mM, and that the peptide:micelle-ratio should be ≤ 0.25 for optimal structuring. Thus, for a 140 µl sample containing 6.8 mg curvacin A, a micelle concentration of about 350 mM DPC would give a peptide:micelle-ratio of 0.25. Hence, for the NMR-sample preparation, 6.8 mg curvacin A was dissolved in 350 mM DPC (see section 3.7). Temperatures between 25 and 35 $^{o}$C were chosen for NMR-experiments, based on the results of the CD-experiments.

Several NMR spectra, both TOCSY and NOESY, of curvacin A were recorded at temperatures of 25, 30 and 35$^{o}$C. The mixing times were 48 ms and 64 ms for the TOCSY spectra, whereas the NOESY spectra had mixing times of 100 ms and 150 ms. By inspecting spectra recorded at different mixing times and temperatures, we found that the most optimal spectra with respect to resolution and number of well-defined peaks were obtained at 35$^{o}$C with a mixing time of 48 ms for TOCSY and 150 ms for NOESY. These were consequently the spectra that were used the most, although additional information was obtained from spectra recorded under different conditions.

There are several protocols that thoroughly describe the processes involved in analysing NMR spectra obtained by different types of NMR techniques [80, 94, 96]. The procedure applied in the structural assignment of curvacin A was described in sections 2.1.7.1 (interpretation of TOCSY spectra) and 2.1.8.1 (interpretation of NOESY spectra). In the following, the analysis of the NMR data obtained on curvacin A is presented.

Curvacin A consists of 41 amino acids, where all but the first (A1) amine proton are potentially visible in the spectra recorded. The amine protons of the first residue in a polypeptide chain cannot readily be observed because the free amine protons are in fast exchange with the protons of the solvent molecules: one proton is exchanged for another in too short a time for a signal to be detected [80, 86]. The remaining 40 amino acids are divided between 12 different spin systems. Table 4.3.1.1 below shows the distribution of amino acids between these spin systems.

Table 4.3.1.1: Distribution of spin systems among the amino acids in curvacin A

| **Spin system** | **Associated amino acids in curvacin A** | **Number of amino acids in curvacin A** |
|---|---|---|
| AMX | S, C, N, Y, W | 13 |
| AM(PT)X | Q, E, M | 3 |
| $A_2(T_2)MPX$ | R | 2 |
| $A_2(F_2T_2)MPX$ | K | 2 |
| $A_3B_3MX$ | V | 2 |
| $A_3MPT(B_3)X$ | I | 3 |
| $A_3B_3MPTX$ | L | 1 |
| $A_3MX$ | T | 1 |
| AX | G | 8 |
| $A_3X$ | A | 4 |
| AA'XX' | Y: ring protons | 2 |
| A(X)MP+A | W: ring protons | 2 |

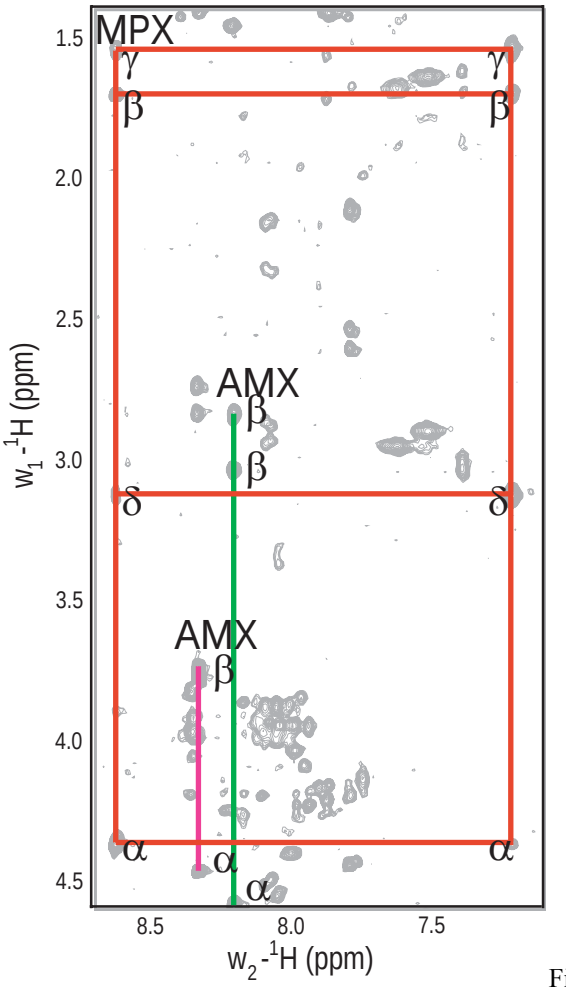


Figure 4.3.1.2: Identification of three spin systems in the (35°C/48 ms) TOCSY spectrum of curvacin A. The peaks are labelled according to the carbon atom with which the hydrogen atoms are bonded. The red line connects a MPX spin system, whereas two AMX spin systems are connected with pink and green lines, respectively. See text for details.

Fig

From the principles described in Figure 2.1.7.1, we can in Figure 4.3.1.2 identify three different spin-systems. The red, pink and green lines connect three different spin systems. The pink AMX system shows a markedly lower $\omega_2$-shift compared to the green system, which is most pronounced for the pair of β-protons. This is typical for a serine residue. The green AMX system is compatible with any of the amino acids associated with an AMX spin system.

The red spin system has unique topology with respect to the two other spin systems that are labelled in Figure 4.3.1.2. The $A_2(T_2)$MPX (MPX for short) can only be one amino acid in curvacin A: arginine (Figure 4.3.1.3). This spin system can therefore be assigned to either R2 or R19.

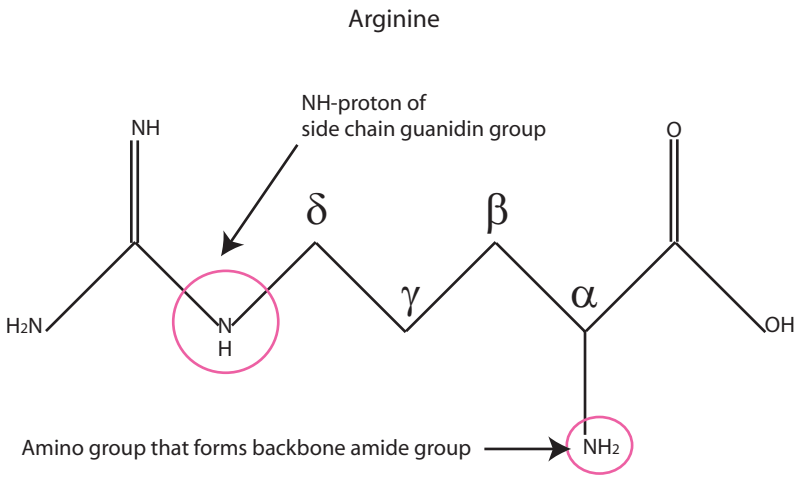


Figure 4.3.1.3: Illustration of the amino acid arginine, highlighting the two amine groups in arginine (purple circles) that correlates to the other side chain protons in Figure 4.3.2.1.

The uniquely recognizable spin system of arginine consists of what appears to be the two parallel spin systems along the $\omega_2$-shift in the amide region. These two parallels are marked with red vertical lines in Figure 4.3.1.2. The horizontal red lines connect the cross peaks between the same side chain protons ($\alpha$, $\beta$, $\gamma$ or $\delta$) and different amine (side chain)/amide (backbone) proton. The only difference between the two parallels is which proton they correlate to. The R-side chain of arginine contains a guanidine group (Figure 4.3.1.3), and thus the one and same spin system correlates to different nitrogen-bonded protons in the amino acid in this region. In this case, correlations from the backbone amide proton (far right of Figure 4.3.1.3) and from the side chain guanidine group-proton in pink circle in Figure 4.3.1.3 (far left of Figure 4.3.1.2) can be seen. Also, the $\delta$-peak around 3 ppm is typical for this type of spin system. This proton is connected to the side chain amide group, and has a much lower chemical shift value than typically observed for other $\delta$-protons.

The lysine residues exhibit the same peak topology as the arginine residues, with an additional side chain proton peak at around 1.5 ppm (lysine has four side chain carbon atoms) (see Figure 2.1.1.1). Lysine has an amino group in the $\zeta$-position of the side chain, and thus the side chain protons correlates to two different amide protons in the same manner as described for arginine. The two lysine residues were therefore easily recognizable from the TOCSY spectrum. Thus, if one locates a cross peak at about 3 ppm that is connected with cross peaks at higher ppm in a TOCSY spectrum, this is indicative of either a lysine or an arginine residue.

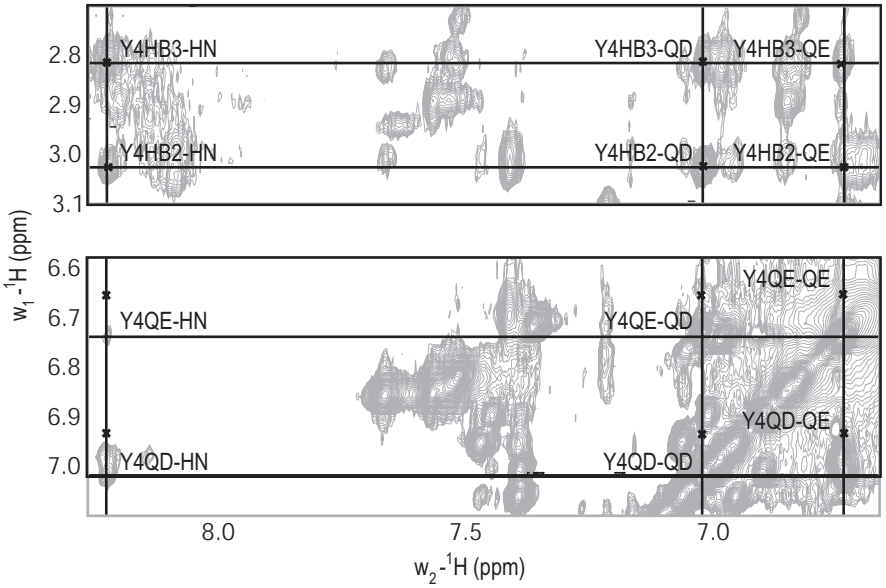


Figure 4.3.1.4 Correlation between the ring protons, amide proton and β-protons of Y4. The peaks are labelled as described in section 2.1.7, with QD and QE representing the (2, 6) and (3, 5) ring proton pairs, respectively. Horizontal and vertical black lines connecting the cross peaks are added for clarity.

The aliphatic side chains of the aromatic residues in curvacin A can be singled out by identifying which AMX-spin systems correlates with the ring protons. The ring protons are positioned in the spectral region around the diagonal at about 6.7-7.5 ppm. Cross peaks between the ring protons and the α- and/or β-protons connects the aliphatic protons and the ring protons in one spin system with unique topology. Locating the aromatic residues in the NOESY spectrum alleviates the sequential assignment procedure described in section 4.4, as there are few aromatic residues in curvacin A. This reduces the number of possible positions of these residues in the sequence, and the proximate residues linked up to the aromatic residues. In Figure 4.3.1.4, this is done for the ring protons, the amide proton and the β-protons of a tyrosine residue (subsequently identified as Y4). The α-proton is left out for clarity as the peak is partly buried in the water peak.

The cross peaks between the side chain protons and the amide proton at the far left of Figure 4.3.1.4 are the same cross peaks as the green AMX-spin system in Figure 4.3.1.2. Thus, the green AMX-spin system must belong to one of the two tyrosine residues (Y4 or Y9).

### 4.3.2 Overlapping peaks and signal-to-noise:

Even though it in theory should be simple to analyse a spectrum, the interpretation of a spectrum may be difficult. The HN-shift values of L38 (7.867) and A39 (7.862) for example are close, and differentiating between the two spin systems can thus be difficult. All the backbone amide protons in curvacin A resonate between 8.7 and 7.6 ppm at 30$^{o}$C. That is 40 spin systems within a range of 1.1 ppm. With this high density of peaks, overlapping peaks complicate the analysis.

The region between 4.3 and 3.7 ppm is very crowded (Figure 4.3.2.1). In this region we find the Hα-HN cross peaks along with the Hβ-HN cross peaks from serine and threonine residues. For example, the α-proton peaks belonging to G40 (coordinates 8.045, 3.955) and G29 (coordinates 8.048, 3.942) are clustered together in what appears to be one single peak. Sequential determination (as described in section 4.4) revealed that two proton peaks (G29-Hα1 and G40-Qα ) were positioned very close to each other (intra residue cross peaks going to both M30 and M41 went from this position, indicating that both G29 and G40 had a proton located at this shift).

The region between 4.3 and 3.7 ppm is in addition close to the resonance of the water protons, which creates noise that complicates the assignment of this part of the spectrum. As can be seen from Figure 4.3.2.1, Y4- and W34 -HA-HN peaks are just on the edge of the ridge, making identification of these peaks difficult.

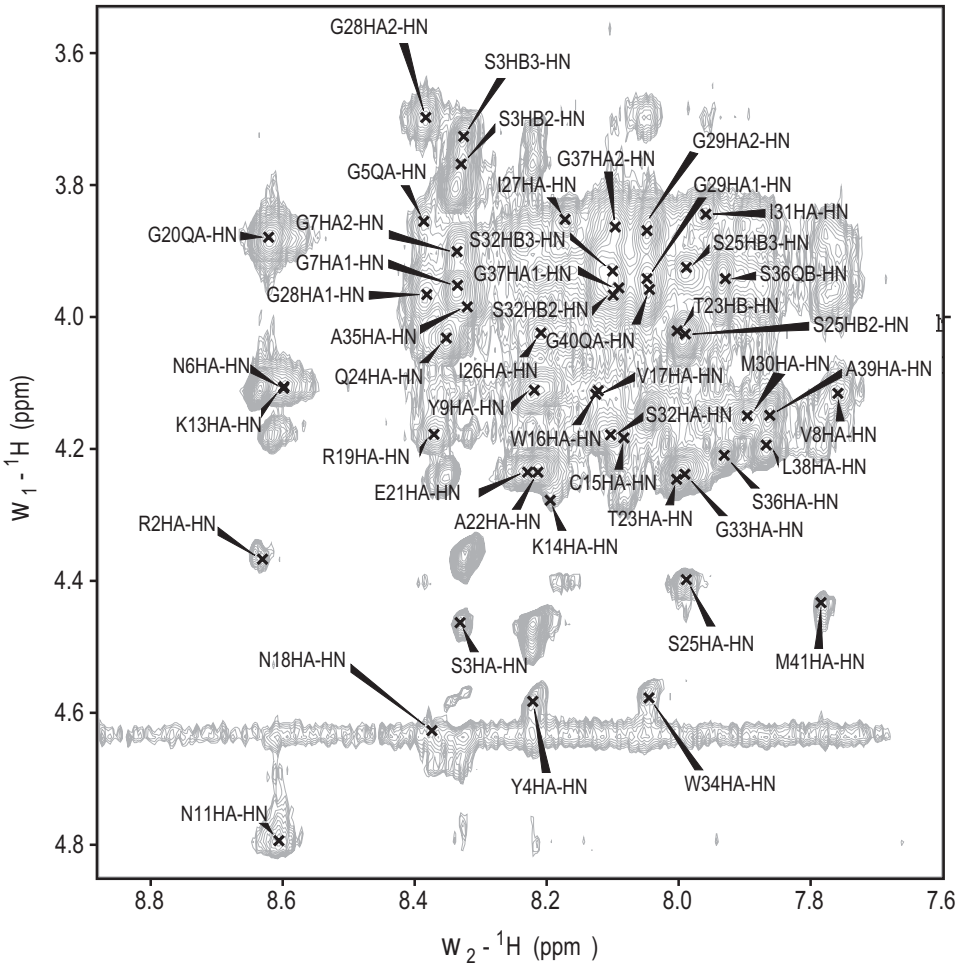


Figure 4.3.2.1 Section of the (150 ms/35°C) NOESY spectrum showing the crowded region with correlation between the α-protons and the amide protons. The peaks are labelled with one-letter code for residue-type and -number according to the position in the sequence, together with the correlated spins giving rise to the cross peak.

Correlations of the same spin systems as in Figure 4.3.2.1 can be found in the alkyl region (see Figure 2.1.7.1 for definition). Here, the correlation from one of the side chain protons in the residue to the other protons is visible. Spin systems that have overlapping HN-shifts may have well-separated e.g Hα-shifts, and one can thus separate two spin systems from each other. By employing all the spectral regions in a NMR-spectrum (fingerprint- amide- and alkyl-region) many ambiguities due to overlap in certain regions can be overcome. The cross peaks in a NMR-spectrum are reflected about the diagonal (see Figure 2.1.8.1), as a result of the direction in which the magnetisation travels. An area with much noise on one side of the diagonal may have a complementary area with less noise on the opposite side of the diagonal. Peaks that are not well defined on one side of the diagonal may thus be well-defined on the opposite side. By tracing complementary peaks on both sides of the diagonal, the identification of these peaks becomes easier.

### 4.3.3 Spectra recorded at different temperatures and with different mixing times

By changing the temperature at which a spectrum is recorded, the position of the water signal changes. The chemical shift of water varies linearly with temperature:

Chemical shift of water (ppm) = 4.766 ppm – (0.0119 ppm $^{o}C^{-1}$)(temperature in $^{o}C$ - 25 $^{o}C$) [87]

The chemical shift of water is thus highly correlated to temperature. A peak that is buried in the water peak at one temperature can be well separated from the water peak at a different temperature. As an example, the N18Hα-peak (coordinates 8.374, 4.627) is buried in the water ridge in the 35$^{o}$C TOCSY spectrum (Figure 4.3.2.1). The peak could be identified by inspecting another spectrum recorded at 25 $^{o}$C, where the water ridge and the N18Hα-peak do not overlap.

The binding of peptide to the micelles is temperature-dependant. There is equilibrium between two different states of the peptide: bound to micelles (structured) and free peptide (unstructured). The interconversion between the two states is in fast exchange, and the population of the unstructured state is low compared with the structured state (CD-results indicate that majority of curvacin A molecules are associated with micelles, and that only a small fraction of peptides are not bound to a micelle (unstructured) at at a given time). As a result only one narrow peak is seen in a spectrum. This peak reflects the average of the two structures (bound and free peptide). The equilibrium conditions are temperature dependant. Thus, by changing the temperature, there is a shift in equilibrium between the two states (the population of the two states changes with respect to each other as a result of the change in the equilibrium conditions) and as a result, the peaks in a spectrum may move. Peaks from different protons that overlapped in one spectrum can be separated at another temperature because the change in temperature affects the equilibriums and causes the two proton peaks to move relative to each other in the spectrum. Changing the temperature does not change which protons correlate with each other. Thus, one can distinguish between two overlapping spin systems by noting which peaks move together when the temperature is changed. For example, in the 35 $^{o}$C spectrum, the amide shift of S3 and N12 overlap at 8.34 ppm. In the 25 $^{o}$C spectrum, however, these two spin systems are well separated at 8.36 (S3) and 8.42 (N12) (Figure 4.3.3.1).

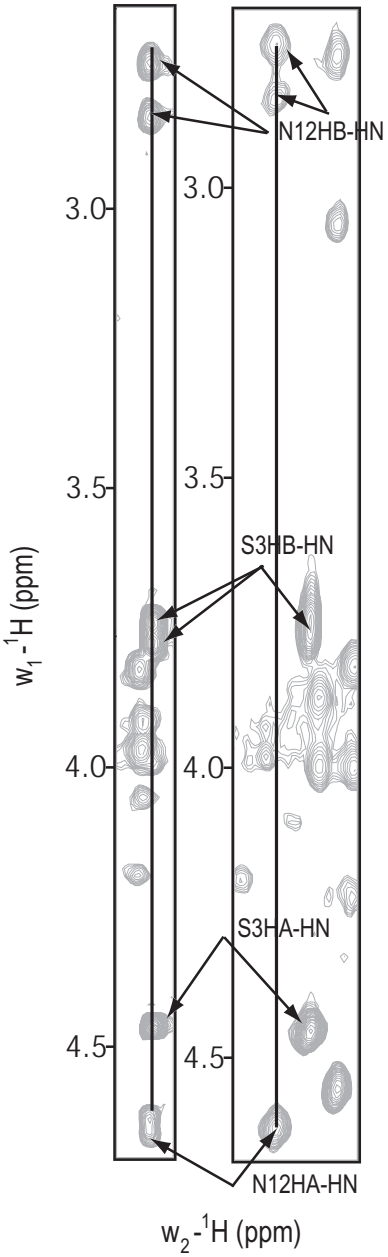


Figure 4.3.3.1 TOCSY spectrum at 35°C/64 ms (left) and TOCSY spectrum at 25 °C/48 ms (right) of the amino acid residues S3 and N12. The N12-spin system is connected with a black vertical line in both spectra. The two spin systems overlap in the $w_2$-dimension in the lefthand (35 °C) spectrum, but are well separated in this dimension in the rigthand (25°C) spectrum.

## 4.4 Sequential assignment and dipolar couplings; NOESY

After having identified the different spin systems in the TOCSY spectrum, the next step was to determine which systems were connected sequentially. This was achieved by analysing a NOESY spectrum. By employing the principles described in Figure 2.1.8.1, three spin systems taken from the NOESY spectrum of curvacin A were identified in Figure 4.4.1.

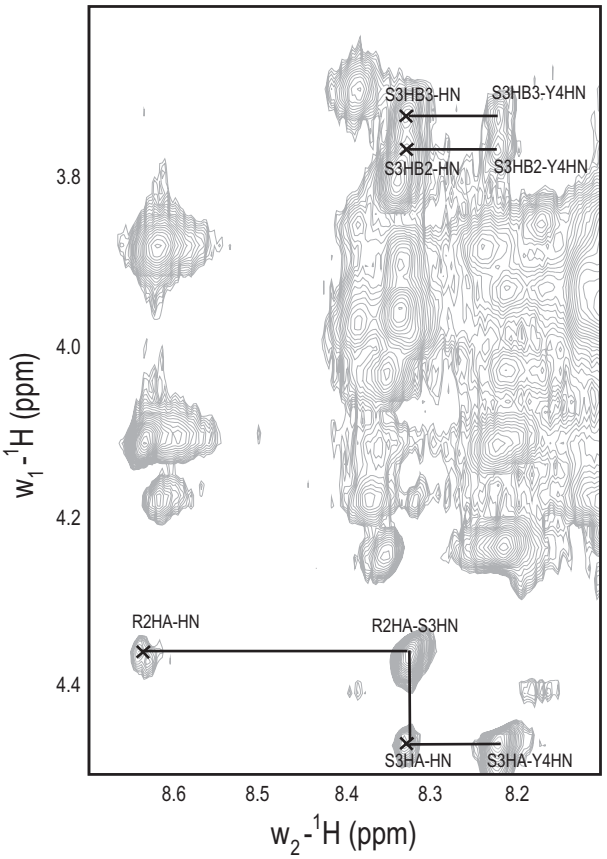


Figure 4.4.1: Connections of three spin systems in the fingerprint region of the (35°C/150 ms) NOESY spectrum. See text for details.

These are the same three spin systems as in Figure 4.3.1.2. They are connected sequentially as described in Figure 2.1.8.1: MPX (residue i) to AMX pink (residue i+1) to AMX green (residue i+2). The next step was to determine where in the sequence these three spin systems are positioned. The interresidual NOE couplings between the spin systems are labelled by stating which proton on a residue contributes to the particular cross peak. For example, the cross peak R2HA-S3HN is the

cross peak between the α-proton of R2 and the amide proton of S3. This is already done in Figure 4.4.1, and the process is described below.

By labelling all the amino acids in the curvacin A sequence according to spin system type, and inspecting where in the sequence these three spin systems occur sequentially, one can determine which amino acids they represent. The MPX spin system can, as described above, only be an arginine. The pink AMX-spin system is most likely a serine, whereas the green AMX-spin system is a tyrosine. Thus the three connected spin systems are R (residue 1), S (residue 2) and Y (residue 3). There is only one place in the sequence where these three spin systems occur sequentially: R2, S3 and Y4.

Assuming that the spin systems were not as easily recognizable, one would have had to determine the sequential arrangement based on the identification of the spin systems. The two AMX spin systems could be either of five amino acids (serine, cysteine, asparagine, tyrosine or tryptophan) (Table 4.3.1.1). There are four possible locations for two consecutive AMX-spin systems in the curvacin A sequence; S3-Y4, Y9-C10, C10-N11 and N11-N12. It was thus not possible to determine which amino acids these two spin systems corresponded to without connecting them up to the next and/or the previous amino acid. The two AMX-spin systems follow sequentially after the MPX-spin system. As we know that the MPX spin system must be an arginine, the three amino acids could be located on one of two stretches along the sequence; R2-S3-Y4 or R19-G20-E21. S3 and Y4 both belong to an AMX-type spin system. G20 belongs to an AX-spin system, whereas E21 belongs to an AM(PT)X spin system. There was thus only one of these stretches of three amino acids in the sequence that fitted the given sequence of spin systems identified in Figure 4.3.1.2, and connected in Figure 4.4.1; R2 (residue i), S3 (residue i+1) and Y4 (residue i+2). The same result is thus obtained from only analysing the type of spin systems as when the identity of the amino acids is known. It is none the less easier to pick out the right sequence if the spin systems are already identified as a particular amino acid. Figure 4.4.1 shows how these three amino acids are connected in the lower part of the fingerprint region. In Figure 4.4.2, supporting cross peaks between the R2 β-protons and S3 amide proton are traced.

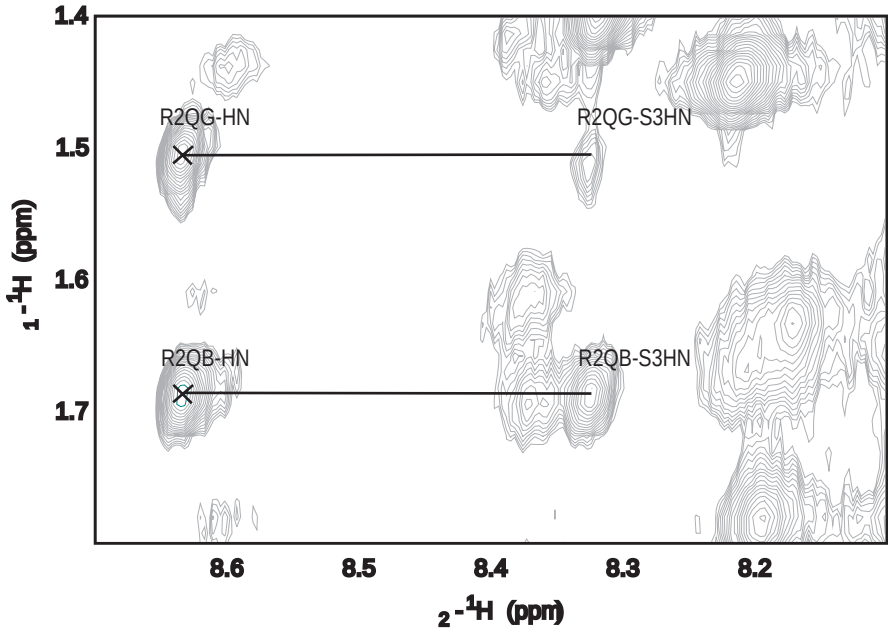


Figure 4.4.2 Cross peaks between R2 and S3 confirming the sequential assignment as determined in Figure 4.4.1.

The sequential connection of the three amino acids is further confirmed by tracing the amide region for cross peaks between backbone amide protons. Ideally, one should be able to locate cross peaks on both sides of the diagonal, as shown in Figure 2.1.8. However, the spectral region close to the diagonal is crowded with peaks that to a large extent overlap and one can often only identify a limited number of these connections with confidence. The NOESY spectrum in Figure 4.4.3 shows the backbone cross peaks for R2-S3-Y4.

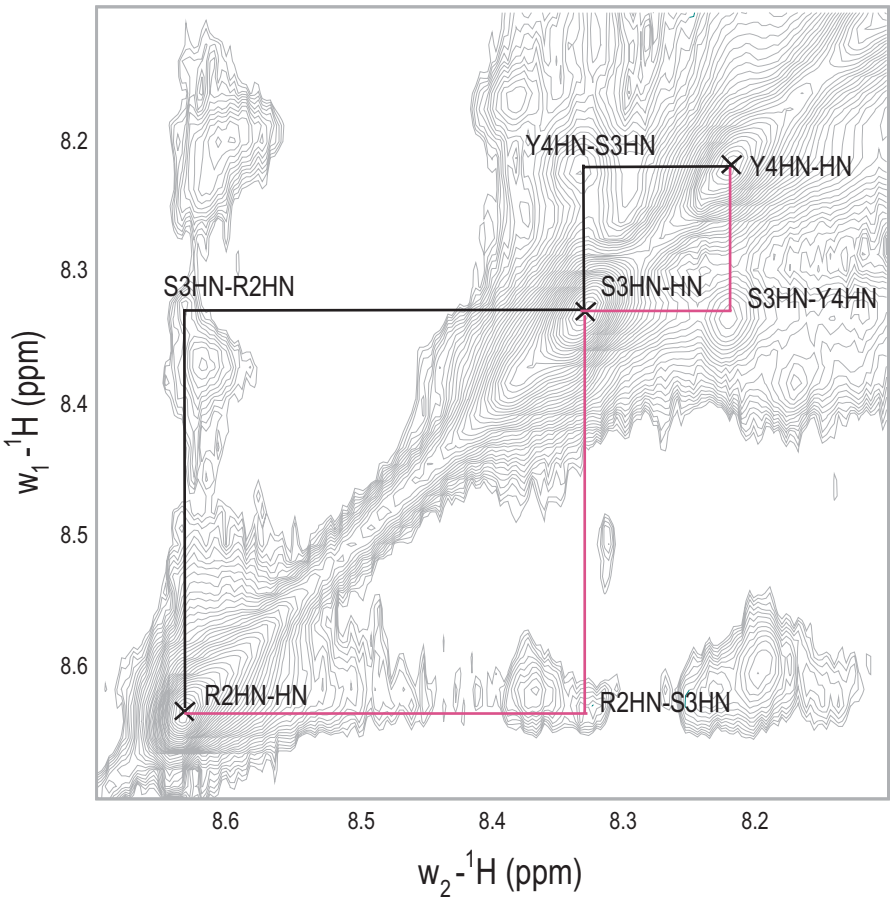


Figure 4.4.3: The three amino acids R2, S3 and Y4 connected by interresidual cross peaks in the amide region of the (35 $^{o}$C/150 ms) NOESY spectrum of curvacin A. The black line shows the connections going from Y4 to R2, whereas the pink line shows the connection going from R2 to Y4.

Cross peaks from R2-Y4 on both sides of the diagonal are marked in Figure 4.4.3. Between S3 and Y4 there is a distinct peak at both side of the diagonal confirming the sequence determined above. The connection between R2 and S3 is clearly identified. The pink line shows the connection from R2 to Y4, whereas the black line shows the opposite direction; from Y4 to R2.

Ideally, connections between all amino acids in a molecule from the first to the last residue should be identified through Hα-HN, Hβ-HN and HN-HN (i, i+1) connections. This is seldom possible, as real spectra are not always easy to assign. Furthermore, the strength of these connections is structure-dependant. In Figure 4.4.4, five consecutive amino acid residues in the Hα-HN region were identified.

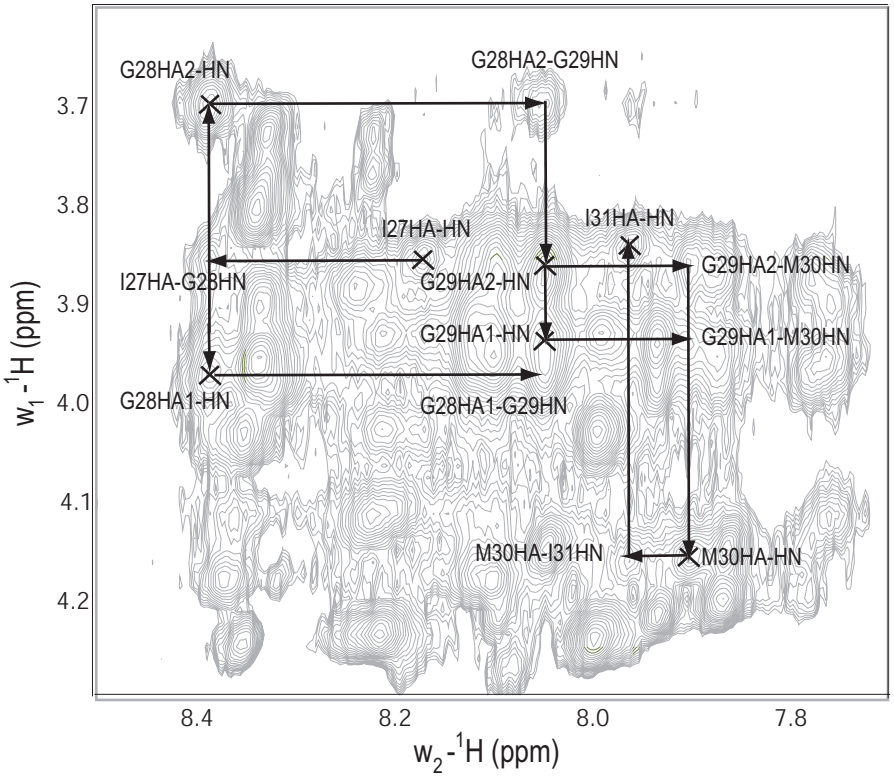


Figure 4.4.4 The amino acid residues I27 through I31 connected by interresidual NOE cross peaks in the (35 °C/ 150 ms) NOESY spectrum. The first NOE cross peak is between α-proton of I27 and the amide proton of G28. The direction of the sequential NOE cross peaks are identified by arrows.

In Table 4.4.1A, the NOEs observed in curvacin A are listed. Sequential connections for all residues were identified through Hα-HN, Hβ-HN and/or HN-HN (i, i+1) connections. All residues are sequentially linked up to each other by at least one (i, i+1) connection. This verifies that the assignment of the spin systems is correct. Multiple connections are especially pronounced in the C-terminal end of curvacin A, where medium range NOEs are also found. The importance of this is discussed in section 4.5 below.

Table 4.4.1 **A** Pattern of interresidue NOEs observed in the NOESY experiment (150 ms mixing time) with curvacin A in the presence of 8 mM DPC. The thickness of the lines shown is relative to the size of the NOE cross peak intensities. The triangles indicate where TALOS found dihedral angle restrains that are typically found for α-helices, stars indicate dihedral angles that can be both α-helix and β-sheet, and circles indicate angles that are found for neither helices nor β-sheets. **B** the CSI indexes [126, 127] obtained by analysis of the Hα and N chemical shift values of curvacin A.

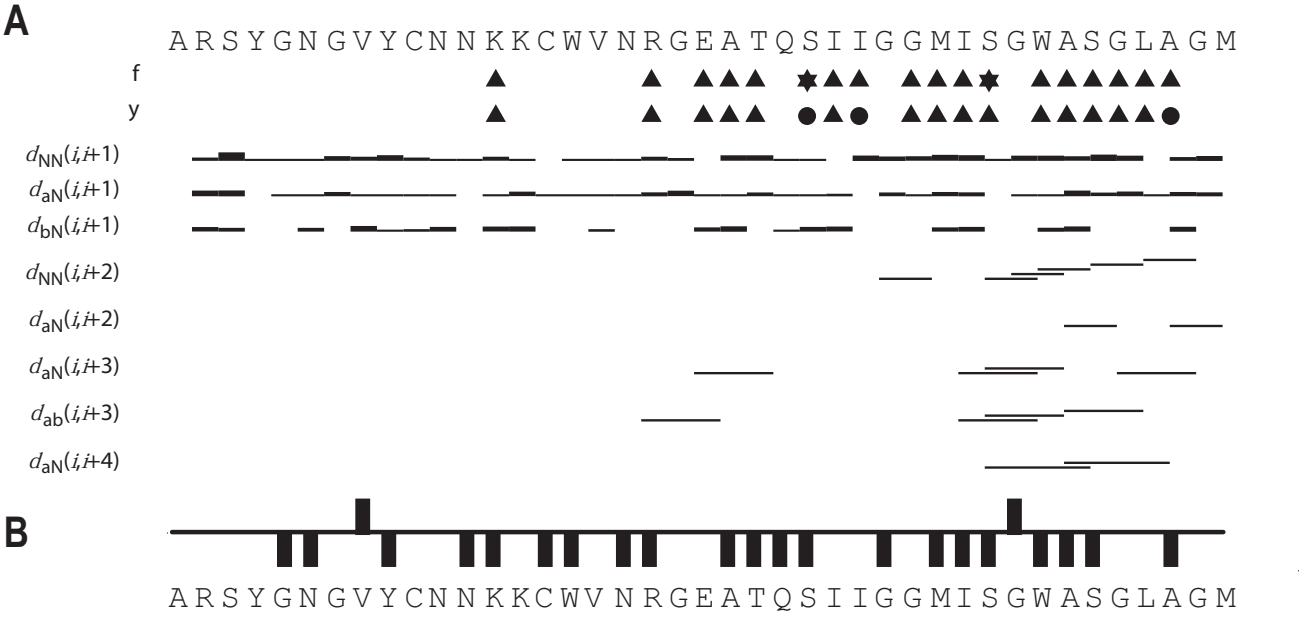


Table 4.4.2 lists the number of NOEs derived from the (35°C/150 ms) NOESY spectrum of curvacin A. There are few long range (R≥5) NOEs. This indicates that the structure of curvacin A is such that it does not allow close contact between residues far apart in the sequence. i.e the structure must be fairly linear.

Table 4.4.2 Number of total NOEs (567) for curvacin A divided into groups according to the range R between the residues. Intraresidual: R=0. Sequential: R=1. Medium range: R<5. Long range: R≥5. Number hydrogen bonds are listed last.

| Number of NOEs and H-bonds | |
|---|---|
| Intraresidual | 323 |
| Long range | 24 |
| Medium range | 78 |
| Sequential | 142 |
| H-bonds | 12 |

## 4.5 Structure prediction of curvacin A and input data for structure calculation

Upper distance constraints are the major input data for structure calculation with Cyana. These distance limits are generated by Cyana based on the list of integrals or intensities of the NOESY peaks and the assignment of these peaks. Upper distance constraints for the known disulfide bridge between C10 and C15 is also included in the calculation. These are based on upper distance limits between disulfide bridges from other predicted structures [128]. From the input data, predicted structural elements of curvacin A are clearly visible. Additional angle constraints from the application of TALOS, and constraints for hydrogen bonds between residues in helical conformations (as predicted by CSI and early structure calculations) are added in the final refinement of the structure calculation. These additional constraints verify and strengthen the already predicted structure of curvacin A from earlier calculations.

CSI and TALOS structure prediction

Information from chemical shift values is of importance in NMR structure calculation, as described in section 2.1.1. In this thesis, secondary structure elements (α-helix, β-sheet or random coil) of curvacin A is predicted by the use of CSI [84] and TALOS [117].

In CSI-prediction, the values -1, 0 and 1 are given to every residue in a peptide segment relative to whether the $^{1}$H-shift is downfield, within or upfield, respectively, of a given threshold value based on the random coil value of that specific residue. Depending on the number of consecutive residues with +1, -1 or a mix of the two, the region is ascribed as having α-helical-, β-sheet- or random coil-conformation (a row of three or more 1 indicates β-sheet, a row of four or more -1 indicates α-helix) The same prediction can be carried out for 13C-shifts, but with the signs for β-sheet and α-helix interconverted [81]. This has proven to be a quick and easy way to confirm already plausible structures, or to hint at as of yet unknown structural elements.

TALOS (Torsion Angle Likelihood Obtained from Shift and sequence similarity) is a program that calculates the protein backbone angle restraints from searching a database for chemical shift and sequence homology [85]. TALOS uses any combination of $^{13}$Cα-, $^{13}$Cβ-, $^{13}$C'-, $^{1}$Hα- and $^{15}$N-shifts tabulated from a given spectrum as input values and gives predicted ϕ– and ψ–angles for each residue. The more shifts added for a residue, the more accurate the output angles calculated.

Currently, a string of three consecutive amino acid residues are used to predict the central residue's angles [85], to ensure that the predicted angles also are in agreement with adjacent peptide bond angles.

From the TALOS prediction presented in Table 4.4.1A, dihedral angles are overall in agreement with an α-helical conformation between residues R19-I27 and G29-A39. The NOE pattern confirms the C-terminal helix, with intraresidual NOE couplings between residues i and/or i+3 and i+4 (which are typical connections for residues in α-helical conformation). There are two i+3 intraresidual NOEs in the region around residues 19-24 indicating a possible central helix. The CSI-prediction in Table 4.5.1B is somewhat more ambiguous, but α-helical regions can be identified between residues A21-S25 and M30-S36 (only one interruption of consecutive -1's). No secondary structure elements could be predicted for the N-terminal part of curvacin A from these results. Thus, from the preliminary structure prediction of curvacin A one can conclude that an unstructured N-terminal region is followed by a possible central helix, and a C-terminal helix.

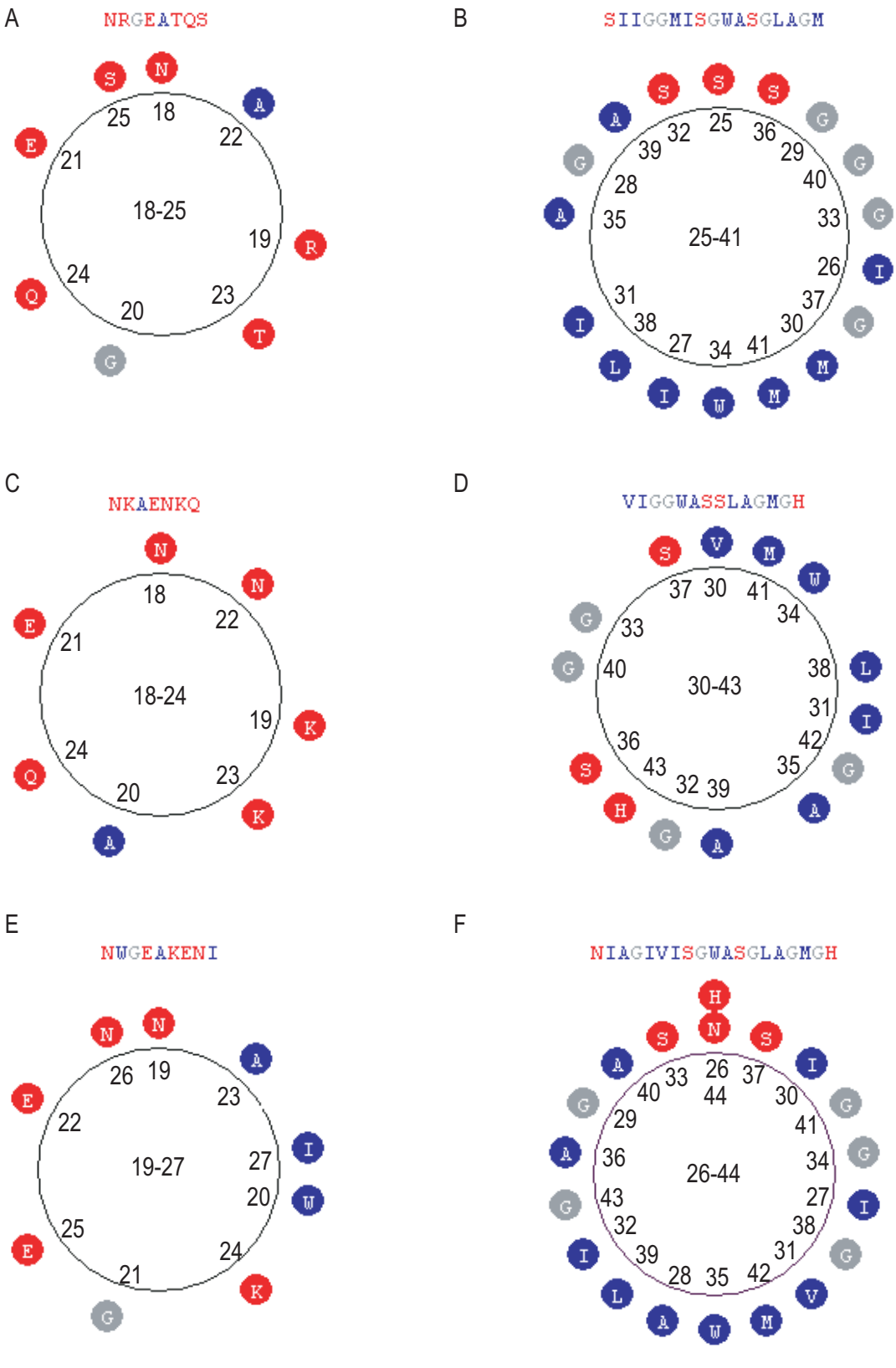
A
NRGEATQS
18-25
B
SIIGGMISGWASGLAGM
25-41
C
NKAENKQ
18-24
D
VIGGWASSLAGMGH
30-43
E
NWGEAKENI
19-27
F
NIAGIVISGWASGLAGMGH
26-44
Hydrophobic
Hydrophilic
Other

Figure 4.4.5 Edmundson helical wheel of fragments of curvacin A (A and B), carnobacteriocin BM1 (C and D) and enterocin P (E and F). Colour-scheme referres to polarity and numbers to residue number in bacteriocin sequence. **A, C and E** helical wheel representation of the central (polar) helix. **B, D and F** helical wheel representation of the amphiphilic C-terminal helix.

Helical wheel analysis

Edmundson helical wheel analysis of curvacin A, carnobacteriocin BM1 and enterocin P (Figure 4.4.5) reveals that all can form a central polar helix and an amphiphilic C-terminal helix. The short central polar helix (residues 18-25 for curvacin A, 18-24 for carnobacteriocin BM1 and 21-26 for enterocin P) is only interrupted by the mildly hydrophobic amino acid alanine (Figure 4.4.5A, C and E). In addition, enterocin P may form a central amphiphilic helix between residues 19 and 27 (Figure 4.4.5E). The amphiphilic C-terminal helix (residues 25-41 for curvacin A, 30-43 for carnobacteriocin BM1 and 26-44 for enterocin P) has a wide hydrophobic side and a narrow hydrophilic side (Figure 4.4.5B, D and F). However, both curvacin A and carnobacteriocin BM1 have a double glycine-motif (G28-G29 for curvacin A, G32-G33 for carnobacteriocin BM1) that is not expected to be in a secondary structural element such as an $\alpha$-helix. The double glycines introduce flexibility to the structure, and therefore an amphiphilic helix might be formed between residues 29-41 (curvacin A) and 33-43 (carnobacteriocin BM1). No such double glycine-motif is found for enterocin P.

## 4.6 Three-dimensional structure of curvacin A as predicted with CYANA

From a total of 567 NOEs, calculations with CYANA and the ANNEAL-routine of CYANA, 20 final structures of curvacin A in DPC micelles were generated. Table 4.6.1 presents some statistical data derived from this calculation.

Table 4.6.1 Statistical data from the structure prediction of curvacin A in DPC-micelles. **A** Deviation from dihedral angles found in secondary structure elements as described in a Ramachandran-plot. **B** The global root mean square distances measured for the calculated ensemble of 20 curvacin A structures by the application of the MOLMOL program [118]

**A**

| Ramachandran deviation (%)* | |
|---|---|
| Most favoured | 63.2 |
| Additionally allowed | 19.5 |
| Generously allowed | 12.9 |
| Disallowed | 4.4 |

**B**

| Global Root-Mean-Square-Difference (RMSD) in Å** | | | | | |
|---|---|---|---|---|---|
| Residues | Backbone*** | All atoms | Residues | Backbone** | All atoms |
| 1-41 | 6.56±1.88 | 7.52±1.95 | 18-24 | 1.09±0.71 | 2.56±1.11 |
| 2-15 | 1.99±0.69 | 2.82±0.79 | 19-24 | 0.63±0.39 | 1.89±0.67 |
| 2-16 | 2.28±0.89 | 3.06±0.96 | 19-25 | 0.70±0.35 | 1.91±0.64 |
| 2-17 | 2.62±1.02 | 3.38±1.10 | 20-24 | 0.28±0.14 | 1.18±0.28 |
| 3-15 | 1.96±0.72 | 2.76±0.83 | 28-39 | 0.37±0.23 | 0.58±0.20 |
| 3-16 | 2.21±0.87 | 2.97±0.97 | 29-39 | 0.18±0.08 | 0.52±0.18 |
| 3-17 | 2.49±0.91 | 3.23±1.13 | 29-40 | 0.23±0.10 | 0.54±0.19+ |
| 19-39 | 2.75±1.03 | 4.37±1.46 | | | |

*Calculated with PROCHECK [129]

**Calculated with MOLMOL[118]

***N, Cα and C

Table 4.6.1A lists the percentage of dihedral angles that are within allowed regions for known secondary structure elements for proteins (α-helix, β-sheet and random coil). Only 4.4% are located in disallowed regions, indicating that the calculated structure is a plausible structure from a stereo chemical point of view. Table 4.6.1B lists RMSD-values for various superimpositions of the structure of curvacin A. A low RMSD with a low uncertainty is consistent with similar structure in the given region for all 20 calculated structures (and therefore a structured region). From these statistics, three structural regions can be identified: the first between residues 2-16, the second between residues 19-24 and the third between residues 29-39. Figure 4.6.1 shows the peptide backbone of the 20 final structures of curvacin A superimposed on these three regions along with a superimposition over the entire molecule.

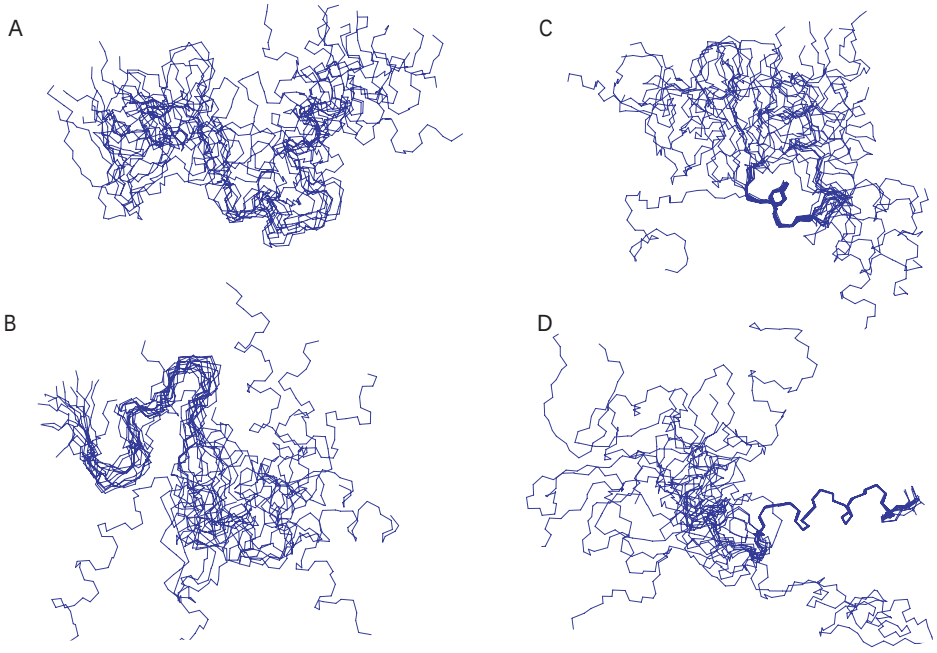


Figure 4.6.1 Superimposition of the 20 final peptide backbone structures of curvacin **A**. A Superimposition over entire molecule (residues 1-41), **B** Superimposition over residues 2-16, **C** Superimposition over residues 19-24 D Superimposition over residues 29-39. Figure prepared with the program MOLMOL [118].

The three structural regions in curvacin A are all clearly observable from Figure 4.6.1: A S-shaped or β-sheet-like domain N-terminally (B), a short central α-helix (C) and a longer C-terminal α-helix (D). The superimposition over the entire molecule (A) shows that the three regions are flexible relative to each other. Evidently, there are two hinge regions in curvacin A: one between the β-sheet-like domain and the central α-helix (residues 16-18) and another between the two helices (residues 25-28). The short C-terminal tail (residues G40 and M41) is unstructured.

The N-terminal domain is less ordered than what has been observed for sakacin P and leucocin A. However, it is more structured than what was observed for carnobacteriocin B2. It was not possible based on the experimental data to assign a perfect β-sheet structure to this domain, although the S-shape indicates a β-sheet-like structure. The β-sheet like structure contains two loops forming an S-

shape. The first loop spans residues 2-10, whereas the second loop runs from residue 10 to 15. Medium range NOEs from R2, S3 and Y4 to V8 and Y9 ties the N-terminal tail up to the second loop of the β-sheet like structure. The cysteine bridge between residues C10 and C15 ties the C-terminal part of this domain covalently together with the starting point of the second loop. The two α-helices are well-defined, as can be seen from Figure 4.6.1. The C-terminal helix is at an angle with respect to the central helix. By close examination of the helices and the side chains of the participating amino acid residues, the central helix was found to be a hydrophilic polar helix (A22 is the only non polar amino acid residue in this helix, and the short side chain of alanine will not contribute markedly with hydrophobicity) (Figure 4.6.2), whereas the C-terminal α-helix was found to be amphiphilic (Figure 4.6.3)


Figure 4.6.2 Peptide backbone representation of curvacin A with the side chains of residues 19-24. **A** Residues 19-24 Residue 20 is a glycine, therefore no side chains are visible for this residue. Polar side chains are coloured red, non polar side chains are coloured green. **B** As in A, but seen down the axis of the helix from R19 to Q24. Figure prepared with the program MOLMOL [118].

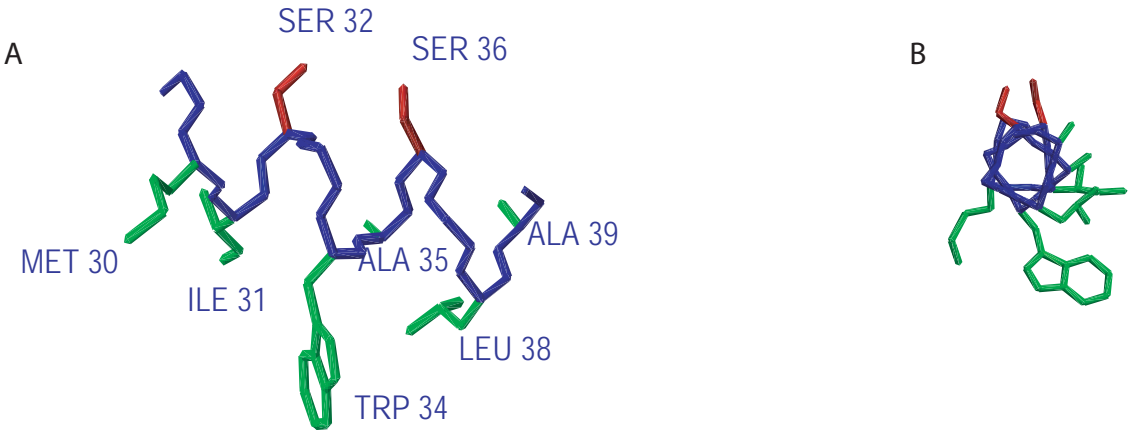


Figure 4.6.3 Peptide backbone representation of curvacin A with side chains of residues 30-39. A Residues 29-39 of curvacin A with non polar side chains in green and polar side chains in red. Residues 29, 33 and 37 are glycines, therefore no side chains are visible for these residues. B As in A, but seen down the axis of the helix from G29 to A39. Polar side chains point up and non polar side chains point down. Figure prepared with the program MOLMOL [118].

The non polar green residues in Figure 4.6.3B all point in the same direction (downwards in the figure), whereas the polar red residues point in the opposite direction. This indicates an amphiphilic helix, with all the polar residues on one side of the helix and all the non polar residues on the other side of the helix. Figure 4.6.4 shows a cartoon depiction of the three-dimensional structure of curvacin A, with the three regions described above.

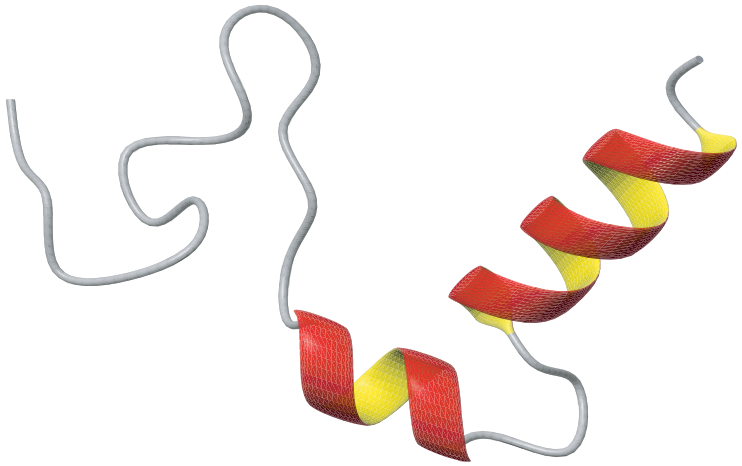

Figure 4.6.4 Cartoon depiction of the three-dimensional structure of curvacin A in DPC-micelles.

## 4.7 Implications of the NMR-structure on the classification of curvacin A and its mode of action

In addition to the structure of curvacin A presented in this thesis, the structure of one other group 3 pediocin-like bacteriocins has been reported: carnobacteriocin B2 [54]. Unlike group 1 (structure solved for sakacin P) and group 2 (structure solved for leucocin A) pediocin-like bacteriocins, sequence alignments between curvacin A and carnobacteriocin B2 does not show large similarities (Figure 2.1.1). The structures of these two group 3 pediocin-like bacteriocins are also substantially different. The structural and sequential differences between curvacin A and carnobacteriocin B2 opens for discussions as to whether or not these two bacteriocins should be placed in the same group [130]. By comparing the amino acid sequence of the group 3 pediocin-like bacteriocins, it might be more relevant to place curvacin A, carnobacteriocin BM1 and enterocin P in one group, and carnobacteriocin B2 and bacteriocin 31 in another, as the three former peptides have similar sequence (especially in the C-terminal half) which is different from the sequence of carnobacteriocin B2 and bacteriocin 31 (Figure 2.1.1).

The three-dimensional structure of the group 3 pediocin-like bacteriocin curvacin A differs from the structures reported for group 1 and 2 pediocin-like bacteriocins. The two helices in curvacin A makes the proposed positioning in the membrane for group 1 and 2 not plausible for curvacin A. An

α-helix has to be approximately 20 amino acid residues long to span a bilayered membrane [131]. The C-terminal α-helix of curvacin A is only 11 residues long and thus cannot span the membrane (as is also the case for group 1 and 2). It is however possible that it penetrates into the membrane-surface interface, and that the C-terminal helix of curvacin A thus has a function equivalent to that of the hairpin-like structure of group 1 and 2 pediocin-like bacteriocins. The polar central helix is not likely to be associated with the hydrophobic part of the membrane, but would rather remain on the outer side of the membrane. The N-terminal domain is also most likely to remain on the outer side of the membrane (possibly associated with a receptor molecule). A possible alternative positioning is presented in Figure 4.7.1. The two helices and the β-sheet like domain are moved relative to each other at the hinges compared with Figure 4.6.4.

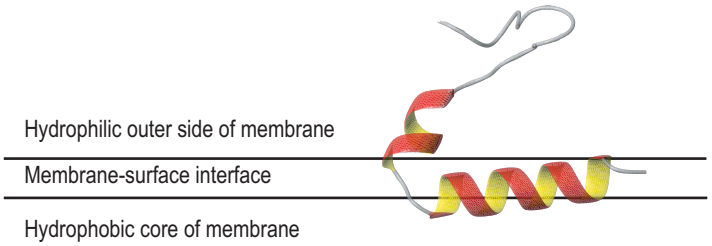


Figure 4.7.1 Possible orientation of curvacin A in a membrane

Activity assays show that the methionines are important for the activity of curvacin A [28]. Oxidation of the methionine residues (M30 and M41), particularly M30, was shown to reduce the activity 10- to 100-fold. This indicates that M30 is important for the activity of curvacin A, and is possibly more deeply buried in the membrane than M41 (not shown in Figure 4.7.1). Solid-state NMR techniques can be applied to verify how curvacin A position itself in the membrane [132-136]. The relative orientation between the peptide and the membrane can be elucidated as the peptide is motionally restricted in solid-state. Fluorescence-studies of the tryptophan-residues may reveal how deep these residues are buried in a membrane [137], which again indicates the molecule's position in the membrane.

*In vitro* site-directed mutagenesis of amino acids in the hinge regions and the two helices may provide valuable information about what kind of conserved residues are important for the activity of curvacin A (conservation of hydrophobicity/hydrophilicity, charged/aromatic residues etc) [57, 107].

From the sequence comparison of different pediocin-like bacteriocins shown in Figure 2.1.1, it is believed that the structure of the curvacin A, enterocin P and carnobacteriocin BM1 have the same overall structure as their sequences are similar.

Finally, it is not only the sequences (and thus possibly the overall structure) of the curvacin A, enterocin P and carnobacteriocin BM1 that are very similar. Also the sequences of the immunity proteins of these bacteriocins are similar, while the immunity protein of carnobacteriocin B2 has quite different sequence [138]. The immunity proteins are thought to specifically recognize the C-terminal domain of the pediocin-like bacteriocins (the hairpin domain of group 1 and 2 pediocin-like bacteriocins) [27], and that the structure of the bacteriocin is important for recognition by the immunity protein. The immunity proteins of curvacin A protect against enterocin P and visa versa [138], indicating that their structure is similar (at least the C-terminal domain that is recognized by the immunity protein). It would be interesting to test whether or not carnobacteriocin BM1 also shares immunity against the other peptides.

# 5 Concluding remarks

The three-dimensional structure of curvacin A in DPC micelles has been determined by the use of NMR-spectroscopy. The molecule has an S-shaped N-terminal domain supported by a disulfide bridge, followed by two helices in the C-terminal half. The central helix is short and polar, whereas the C-terminal helix is longer and amphiphilic. There is a hinge between the S-shaped N-terminal domain and the central helix and a hinge between the central helix and the C-terminal helix, and this enables the three regions to move relative to each other The helix-hinge-helix structure found in the C-terminal half of curvacin A is quite different from the hairpin-like structure found for sakacin P (group 1) and leucocin A (group 2), indicating that the group 3 pediocin-like bacteriocins might have a somewhat different mode of action and position in the membrane than what has been found for group 1 and 2. The C-terminal structure of the group 3 bacteriocins enterocin P and carnobacteriocin BM1 is believed to be equivalent to the one observed for curvacin A, based on sequence similarities.

The structure of curvacin A differs markedly from the structure found for carnobacteriocin B2 (another group 3 pediocin-like bacteriocin). On the basis of this and their amino acid sequence, it might be relevant to separate the group 3 bacteriocins curvacin A, enterocin P and carnobacteriocin BM1 from the rest of the bacteriocins in this group. Other, as of now, unknown structures of group 3 pediocin-like bacteriocins have to be elucidated to verify this new grouping.

The structure of curvacin A may be obtained from the protein data bank (PDB) [129]. The PDB ID code is 2A2B.

# 6 Reference list


1. Ennahar, S., et al., *Class IIa Bacteriocins: Biosynthesis, Structure and Activity.* FEMS Microbiology Reviews, 2000. **24**: p. 85-106.
2. Salminen, S., A.v. Wright, and A. Ouwehand, eds. *Lactic Acid Bacteria: Microbiological and Functional Aspects*. 3. Edition, Revised and Expanded ed. Food Science and Technology, A Series of Monographs, Textbooks and Reference Books. 2004, Marcel Dekker Inc.: New York. 633.
3. Eijsink, V.G.H., et al., *Production of Class II Bacteriocins by Lactic Acid Bacteria; an Example of Biological Warfare and communication.* Antonie van Leeuwenhoek, 2002. **81**: p. 639-654.
4. Jack, R., J.R. Tagg, and B. Ray, *Bacteriocins of Gram-Positive Bacteria.* Microbiological Reviews, 1995. **59**: p. 171-200.
5. Riley, M.A. and J.E. Wertz, *Bacteriocins: Evolution, Ecology and Application.* Annual Review of Microbiology, 2002. **56**: p. 117-1137.
6. Cleveland, J., et al., *Bacteriocins: Safe, Natural Antimicrobials for Food Preservation.* International Journal of Food Microbiology, 2001. **71**: p. 1-20.
7. Ming, X.T., et al., *Bacteriocins Applied to Food Packaging Materials to Inhibit Listeria Monocytogenes on Meats.* Journal of Food Science, 1997. **62**(2): p. 413-415.
8. Abee, T., L. Krockel, and C. Hill, *Bacteriocins: Modes of Action and Potentials in Food Preservation and Control of Food Poisoning.* International Journal of Food Microbiology, 1995. **28**: p. 169-185.
9. Nissen-Meyer, J. and I.F. Nes, *Ribosomally Syntheszied Antimicrobial Peptides: Their Function, Structure, Biogenesis, and Mechanism of Action.* Archives of Microbiology, 1997. **167**: p. 67-77.
10. Sahl, H.G., R.W. Jack, and G. Bierbaum, *Biosynthesis and biological activities of lantibiotics with unique post-translational modifications.* Eur J Biochem, 1995. **230**(3): p. 827-53.
11. Nes, I.F. and J.R. Tagg, *Novel lantibiotics and their pre-peptides.* Antonie Van Leeuwenhoek, 1996. **69**(2): p. 89-97.
12. Miller, K.W., et al., *Isolation and characterization of pediocin AcH chimeric protein mutants with altered bactericidal activity.* Appl Environ Microbiol, 1998. **64**(6): p. 1997-2005.
13. Nieto Lozano, J.C., et al., *Purification and amino acid sequence of a bacteriocin produced by Pediococcus acidilactici.* J Gen Microbiol, 1992. **138**(9): p. 1985-90.
14. Eijsink, V.G.H., et al., *Comparative Studies of Class IIa Bacteriocins of Lactic Acid Bacteria.* Applied Microbiology and Biotechnology, 1998. **64**: p. 3275-3281.
15. Henderson, J.T., A.L. Chopko, and P.D. van Wassenaar, *Purification and primary structure of pediocin PA-1 produced by Pediococcus acidilactici PAC-1.0.* Arch Biochem Biophys, 1992. **295**(1): p. 5-12.
16. Uteng, M., et al., *Three-Dimensional Structure in Lipid Micelles of the Pediocin-like Antimicrobial Peptide Sakacin P and a Sakacin P Variant That Is Structurally Stabilized by an Inserted C-Terminal Disulfide Bridge.* Biochemistry, 2003. **42**: p. 11417-11426.
17. 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.* Mol Microbiol, 1995. **17**(3): p. 515-22.

18. Quadri, L.E., et al., *Characterization of a locus from Carnobacterium piscicola LV17B involved in bacteriocin production and immunity: evidence for global inducer-mediated transcriptional regulation.* J Bacteriol, 1997. **179**(19): p. 6163-71.
19. Nes, I.F., et al., *Unmodified peptide-bacteriocins (class II) produced by lactic acid bacteria*, in *Peptide antibiotics: discovery, modes of action and application*, C.J. Dutton, et al., Editors. 2002, Marcel Dekker: New York. p. 81-115.
20. Marugg, J.D., et al., *Cloning, expression, and nucleotide sequence of genes involved in production of pediocin PA-1, and bacteriocin from Pediococcus acidilactici PAC1.0.* Appl Environ Microbiol, 1992. **58**(8): p. 2360-7.
21. 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, 1996. **142 ( Pt 6)**: p. 1437-48.
22. Axelsson, L. and A. Holck, *The genes involved in production of and immunity to sakacin A, a bacteriocin from Lactobacillus sake Lb706.* J Bacteriol, 1995. **177**(8): p. 2125-37.
23. 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.
24. Havarstein, L.S., H. Holo, and I.F. Nes, *The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by gram-positive bacteria.* Microbiology, 1994. **140 ( Pt 9)**: p. 2383-9.
25. Nes, I.F. and V. Eijsink, *Regulation of group II peptide bacteriocin synthesis by quorum-sensing mechanisms.*, in *Cell-cell signaling in bacteria*, G.M. Dunny and S.C. Winans, Editors. 1999, American Society for Microbiology: Washington. p. 175-192.
26. 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. 1579-96.
27. Fimland, G., L. Johnsen, and B. Dalhus, *The Pediocin-Like Antimicrobial Peptides (Class IIa Bacteriocins) and Their Immunity Proteins: Biosynthesis, Structure, and Mode of Action.* Submitted for publication, 2005.
28. Fimland, G., et al., *New Biologically Active Hybrid Bacteriocins Constructed by Combining Regions from Various Pediocin-Like Bacteriocins: the C-Terminal Region Is Important for Determening Specificity.* Applied and Environmental Microbiology, 1996. **62**: p. 3313-3318.
29. Aymerich, T., et al., *Biochemical and genetic characterization of enterocin A from Enterococcus faecium, a new antilisterial bacteriocin in the pediocin family of bacteriocins.* Appl Environ Microbiol, 1996. **62**(5): p. 1676-82.
30. Metivier, A., et al., *Divercin V41, a new bacteriocin with two disulphide bonds produced by Carnobacterium divergens V41: primary structure and genomic organization.* Microbiology, 1998. **144 ( Pt 10)**: p. 2837-44.
31. Tahiri, I., et al., *Purification, characterization and amino acid sequencing of divergicin M35: a novel class IIa bacteriocin produced by Carnobacterium divergens M35.* Int J Food Microbiol, 2004. **97**(2): p. 123-36.
32. Le Marrec, C., et al., *Biochemical and genetic characterization of coagulin, a new antilisterial bacteriocin in the pediocin family of bacteriocins, produced by Bacillus coagulans I(4).* Appl Environ Microbiol, 2000. **66**(12): p. 5213-20.
33. Yamazaki, K., et al., *Purification and characterization of a novel class IIa bacteriocin, piscicocin CS526, from surimi-associated Carnobacterium piscicola CS526.* Appl Environ Microbiol, 2005. **71**(1): p. 554-7.

34. Tichaczek, P.S., R.F. Vogel, and W.P. Hammes, *Cloning and sequencing of sakP encoding sakacin P, the bacteriocin produced by Lactobacillus sake LTH 673.* Microbiology, 1994. **140 ( Pt 2)**: p. 361-7.
35. Kalmokoff, M.L., et al., *Identification of a new plasmid-encoded sec-dependent bacteriocin produced by Listeria innocua 743.* Appl Environ Microbiol, 2001. **67**(9): p. 4041-7.
36. Bennik, M.H.J., et al., *A Novel Bacteriocin With a YGNGV Motif from Vegetalbe-Associated Enterococcus Mundtii; Full Characterizationand Interaction with Target Organisms.* Biochimica et biophysica acta, 1998. **1373**: p. 47-58.
37. Kawamoto, S., et al., *Biochemical and Genetic Characterization of Mundticin KS, an Antilisterial Peptide Produced by Enterococcus Mundtii NFRI 7393.* Applied and Environmental Microbiology, 2002. **68**: p. 3830-3840.
38. Jack, R.W., et al., *Characterization of the chemical and antimicrobial properties of piscicolin 126, a bacteriocin produced by Carnobacterium piscicola JG126.* Appl Environ Microbiol, 1996. **62**(8): p. 2897-903.
39. Bhugaloo-Vial, P., et al., *Purification and Amino Acid Sequence of Piscicocins V1a and V1b that Display Significantly Different Levels of Specific Inhibitory Activity.* Applied and Environmental Microbiology, 1996. **62**: p. 4410-7500.
40. Vaughan, A., et al., *An analysis of bacteriocins produced by lactic acid bacteria isolated from malted barley.* J Appl Microbiol, 2001. **91**(1): p. 131-8.
41. Fimland, G., K. Sletten, and J. Nissen-Meyer, *The complete amino acid sequence of the pediocin-like antimicrobial peptide leucocin C.* Biochem Biophys Res Commun, 2002. **295**(4): p. 826-7.
42. Hastings, J.W., et al., *Characterization of leucocin A-UAL 187 and cloning of the bacteriocin gene from Leuconostoc gelidum.* J Bacteriol, 1991. **173**(23): p. 7491-500.
43. Fleury, Y., et al., *Covalent Structure, Synthesis, and Structure-Function Studies of Mesentericin Y 10537, a Defensive Peptide from Gram-Positive Bacteria Leconostoc Mesenteroides.* Journal of Biological Chemistry, 1996. **271**: p. 14421-14429.
44. Ferchichi, M., et al., *Lactococcin MMFII, a novel class IIa bacteriocin produced by Lactococcus lactis MMFII, isolated from a Tunisian dairy product.* FEMS Microbiol Lett, 2001. **205**(1): p. 49-55.
45. Simon, L., et al., *Sakacin g, a new type of antilisterial bacteriocin.* Appl Environ Microbiol, 2002. **68**(12): p. 6416-20.
46. Van Reenen, C.A., et al., *Characterization and heterologous expression of a class IIa bacteriocin, plantaricin 423 from Lactobacillus plantarum 423, in Saccharomyces cerevisiae.* Int J Food Microbiol, 2003. **81**(1): p. 29-40.
47. Atrih, A., et al., *Mode of action, purification and amino acid sequence of plantaricin C19, an anti-Listeria bacteriocin produced by Lactobacillus plantarum C19.* Int J Food Microbiol, 2001. **68**(1-2): p. 93-104.
48. Tichaczek, P.S., R.F. Vogel, and W.P. Hammes, *Cloning and Sequencing of curA Encoding Curvacin A, the Bacteriocin Produced by Lactobacillus Curvatus LTH1174.* Archives of Microbiology, 1993. **160**(4): p. 279-283.
49. Holck, A., et al., *Purification and amino acid sequence of sakacin A, a bacteriocin from Lactobacillus sake Lb706.* J Gen Microbiol, 1992. **138**(12): p. 2715-20.
50. Quadri, L.E., et al., *Chemical and genetic characterization of bacteriocins produced by Carnobacterium piscicola LV17B.* J Biol Chem, 1994. **269**(16): p. 12204-11.
51. Cintas, L.M., et al., *Biochemical and genetic characterization of enterocin P, a novel sec-dependent bacteriocin from Enterococcus faecium P13 with a broad antimicrobial spectrum.* Appl Environ Microbiol, 1997. **63**(11): p. 4321-30.

52. Tomita, H., et al., *Cloning and genetic organization of the bacteriocin 31 determinant encoded on the Enterococcus faecalis pheromone-responsive conjugative plasmid pYI17.* J Bacteriol, 1996. **178**(12): p. 3585-93.
53. Fregeau Gallagher, N.L., et al., *Three-Dimensional Structure of Leucocin A in Trifluoroethanol and Dodecylphosphocholine Micelles: Spatial Location of Residues Critical for Biological Activity in Type IIa Bacteriocins from Lactic Acid Bacteria.* Biochemistry, 1997. **36**: p. 15062-15072.
54. Wang, Y., et al., *Solution Structure of Carnobacteriocin B2 and Implications for Structure-Activity Relationships among Type IIa Bacteriocins from Lactic Acid Bacteria.* Biochemistry, 1999. **38**: p. 15438-15447.
55. Watson, R.M., et al., *Conformational changes in pediocin AcH upon vesicle binding and approximation of the membrane-bound structure in detergent micelles.* Biochemistry, 2001. **40**(46): p. 14037-46.
56. Roccatano, D., et al., *Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: a molecular dynamics study.* Proc Natl Acad Sci U S A, 2002. **99**(19): p. 12179-84.
57. Fimland, G., V.G.H. Eijsink, and J. Nissen-Meyer, *Mutational Analysis of the Role of Tryptophan Residues in an Antimicrobial Peptide.* Biochemistry, 2002. **41**: p. 9508-9515.
58. Johnsen, L., G. Fimland, and J. Nissen-Meyer, *The C-terminal domain of pediocin-like antimicrobial peptides (class IIa bacteriocins) is involved in specific recognition of the C-terminal part of cognate immunity proteins and in determining the antimicrobial spectrum.* J Biol Chem, 2005. **280**(10): p. 9243-50.
59. Fimland, G., et al., *A C-Terminal Disulfide Bridge in Pediocin-Like Bacteriocins Renders Bacteriocin Activity Less Temperature Dependent and Is a Major Determinant of the Antimicrobial Spectrum.* Journal of Bacteriology, 2000. **182**: p. 2643-2648.
60. Dalet, K., et al., *σ(54)-dependent PTS permease of the mannose family is responsible for sensitivity of Listeria monocytogenes to mesentericin Y105.* Microbiology, 2001. **147**: p. 3263-3269.
61. Gravesen, A., et al., *High-level resistance to class IIa bacteriocins is associated with one general mechanism in Listeria monocytogenes.* Microbiology, 2002. **148**(Pt 8): p. 2361-9.
62. Gravesen, A., et al., *Restriction fragment differential display of pediocin-resistant Listeria monocytogenes 412 mutants shows consistent overexpression of a putative beta-glucoside-specific PTS system.* Microbiology, 2000. **146 ( Pt 6)**: p. 1381-9.
63. Ramnath, M., et al., *Expression of mptC of Listeria monocytogenes induces sensitivity to class IIa bacteriocins in Lactococcus lactis.* Microbiology, 2004. **150**(Pt 8): p. 2663-8.
64. Ramnath, M., et al., *Absence of a putative mannose-specific phosphotransferase system enzyme IIAB component in a leucocin A-resistant strain of Listeria monocytogenes, as shown by two-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis.* Appl Environ Microbiol, 2000. **66**(7): p. 3098-101.
65. Hechard, Y. and H.G. Sahl, *Mode of action of modified and unmodified bacteriocins from Gram-positive bacteria.* Biochimie, 2002. **84**(5-6): p. 545-57.
66. Chen, Y., R.D. Ludescher, and T.J. Montville, *Electrostatic Interactions, but not the YNGNV Consensus Motif, Govern the Binding of Pediocin PA-1 and its Fragments of Phospholipid Vesicles.* Applied and Environmental Microbiology, 1997. **63**: p. 4770-4777.
67. Kazazic, M., J. Nissen-Meyer, and G. Fimland, *Mutational analysis of the Role of Charged Residues in Target-Cell Binding, Potency and Specificity of the Pediocin-Like Bacteriocin Sakacin P.* Microbiology, 2002. **148**: p. 2019-2027.

68. Miller, K.W., et al., *Production of active chimeric pediocin AcH in Escherichia coli in the absence of processing and secretion genes from the Pediococcus pap operon.* Appl Environ Microbiol, 1998. **64**(1): p. 14-20.
69. Chikindas, M.L., et al., *Pediocin PA-1, a bacteriocin from Pediococcus acidilactici PAC1.0, forms hydrophilic pores in the cytoplasmic membrane of target cells.* Appl Environ Microbiol, 1993. **59**(11): p. 3577-84.
70. Montville, T.J. and Y. Chen, *Mechanistic Action of Pediocin and Nisin: Recent Progress and Unresolved Questions.* Applied Microbiology and Biotechnology, 1998. **50**: p. 511-519.
71. Chen, Y., et al., *Functional Characterization of Pediocin PA-1 Binding to Liposomes in the Absence of a Protein Receptor and its Relationship to a Predicted Tertiary Structure.* Applied and Environmental Microbiology, 1997. **63**: p. 524-531.
72. Killian, J.A. and G. von Heijne, *How proteins adapt to a membrane-water interface.* Trends Biochem Sci, 2000. **25**(9): p. 429-34.
73. Purcell, E.M., H.C. Torrey, and R.V. Pound, *Resonance Absorbtion by Nuclear Magnetic Moments in a Solid.* Physical Review, 1946. **69**: p. 37-38.
74. Bloch, F., W.W. Hansen, and M. Packard, *Nuclear Induction.* Physical Review, 1946. **69**: p. 127.
75. Laider, K.J., J.H. Meiser, and B.C. Sanctuary, *Physical Chemistry*. 4th. edition ed. 2002: Houghton Mifflin Company. 1060.
76. Atkins, P. and J. de Paula, *Atkins' Physical Chemistry*. 7 ed. 2002: Oxford University Press. 1149.
77. Rattle, H., *An NMR Primer for Life Scientists*. 1995: Partnership Press. 124.
78. Claridge, T.D.W., *High-Resolution NMR Techniques in Organic Chemistry*. Tetrahedron Organic Chemistry Series, ed. J.E. Baldwin, F. Williams, and R.M. Williams. Vol. 19. 1999. 381.
79. Friebolin, H., *Basic One- and Two-Dimensional NMR Spectroscopy*. Second, enlarged edition ed. 1993: VCH Verlagsgesellschaft. 368.
80. Wütrich, K., *NMR of Proteins and Nucleic Acids*. 1986: John Wiley & Sons. 292.
81. Wishart, D.S. and B.D. Sykes, *The $^{13}C$ Chemical-Shift Index: A Simple Method for the Identification of Protein Secondary Structure Using $^{13}C$ Chemical-Shift Data.* Journal of Biomolecular NMR, 1994. **4**: p. 171-180.
82. Wishart, D.S. and B.D. Sykes, *Chemical Shift as a Tool for Structure Determination.* Methods in Enzymology, 1994. **239**: p. 363-392.
83. 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**: p. 311-333.
84. 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**: p. 1647-1651.
85. Cornilescu, G., F. Delaglio, and A. Bax, *Protein Backbone Angle Restraints from Searching a Database for Chemical Shift and Sequence Homology.* Journal of Biomolecular NMR, 1999. **13**: p. 289-302.
86. Cavanagh, J., et al., *Protein NMR Spectroscopy. Principles and Practice*. 1996: Academic Press. 587.
87. Wishart, D.S., et al., *$^{1}H$, $^{13}C$ and $^{15}N$ Chemical Shift Referencing in biomolecular NMR.* Journal of Biomolecular NMR, 1995. **6**: p. 135-140.

88. Wishart, D.S., et al., *$^{1}H$, $^{13}C$ and $^{15}N$ Random Coil NMR Chemical Shifts of the Common Amino Acids. I. Investigation of Nearest-Neighbour Effects.* Journal of Biomolecular NMR, 1995. **5**: p. 67-81.
89. Spera, S. and A. Bax, *Empirical Correlation Between Protein Backbone Conformation and C.alpha. and C.beta. 13C Nuclear Magnetic Resonance Chemical Shifts.* Journal Of The American Chemical Society, 1991. **113**(14): p. 5490-5492.
90. Ernst, R.R., G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. The International Series of Monographs on Chemistry, ed. M.L.H. Green, et al. Vol. 14. 1987: Oxford University Press. 610.
91. Jeener, J. in *Ampére Summer School*. 1971. Basko Polje, Yugoslavia.
92. Aue, W.P., E. Bartholdi, and R.R. Ernst, *Two-dimensional Spectroscopy. Application to Nuclear Magnetic Resonance.* Journal of Chemical Physics, 1976. **64**(5): p. 2229-2246.
93. Braunschweiler, L. and R.R. Ernst, *Coherence Transfer by Isotropic Mixing: Application to Proton Correlation Spectroscopy.* Journal of Magnetic Resonance, 1983. **53**(3): p. 521-528.
94. Redfield, C., *Resonance Assignment Strategies for Small Proteins*, in *NMR of Macromolecules. A Practical Approach*, G.C.K. Roberts, Editor. 1993, Oxford University Press. p. 71-99.
95. Levitt, M.H., *Spin Dynamics: Basics of Nuclear Magnetic Resonance*. 2001: John Wiley & Sons, Ltd. 686.
96. Markley, J.L. and M. Kainosho, *Stable Isotope labelling and Resonance Assignments in Larger Proteins*, in *NMR of Macromolecules. A Practical Approach*, G.C.K. Roberts, Editor. 1993, Oxford University Press. p. 101-148.
97. Bax, A. and S. Grzesiek, *Methodological Advances in Protein NMR*, in *NMR of Proteins*, G.M. Clore and A.M. Gronenborn, Editors. 1993, MacMillan Press Ltd. p. 307.
98. Jee, J. and P. Güntert, *Influence of the Completeness of Chemical Shift Assignments on NMR Structures Obtained with Automated NOE Assignment.* Journal of Structural and Functional Genomics, 2003. **4**: p. 179-189.
99. Güntert, P., C. Mumenthaler, and K. Wütrich, *Torsion Angle Dynamics for NMR Structure Calculation with the New Program DYANA.* Journal of Molecular Biology, 1997. **273**: p. 283-298.
100. Güntert, P., *Automated NMR Structure Calculation With CYANA*, in *Methods in Molecular Biology*, A.K. Downing, Editor. 2004, Humana Press Inc.: Totowa. p. 353-378.
101. Herrmann, T., P. Güntert, and K. Wütrich, *Protein NMR Structure Determination with Automated NOE Assignment Using the New Software CANDID and the Torsion Angle Dynamics Algorithm DYANA.* Journal of Molecular Biology, 2002. **319**: p. 209-227.
102. Sheehan, D., *Physical Biochemistry: Principles and Applications*. 2000: John Wiley and Sons, Ltd. 349.
103. Hillenkamp, F. and M. Karas, *Mass Spectrometry of Peptides and Proteins by Matrix-Assisted Ultraviolet Laser Desorption/Ionization.* Methods in Enzymology, 1990. **193**: p. 280-295.
104. Beavis, R.C. and B.T. Chait, *Matrix-Assisted Laser Desorption Ionization Mass-Spectrometry of Proteins.* Methods in Enzymology, 1996. **270**: p. 519-551.
105. Karas, M. and F. Hillenkamp, *Laser Desorption Ionization of Proteins With Molecular Masses Exceeding 10,000 Daltons.* Analytical Chemistry, 1988. **60**(20): p. 2299-2301.
106. Sletten, K. and G. Husby, *The complete amino-acid sequence of non-immunoglobulin amyloid fibril protein AS in rheumatoid arthritis.* Eur J Biochem, 1974. **41**(1): p. 117-25.
107. Johnsen, L., et al., *Engineering Increased Stability in the Antimicrobial Peptide Pediocin PA-1.* Applied and Environmental Microbiology, 2000. **66**: p. 4798-4802.

108. Axelsson, L., et al., *A system for heterologous expression of bacteriocins in Lactobacillus sake.* FEMS Microbiol Lett, 1998. **168**(1): p. 137-43.
109. Uteng, M., et al., *Rapid Two-Step Procedure for Large-Scale Purification of Pediocin-Like Bacteriocins and Other Cationic Antimicrobial Peptides from Complex Culture Medium.* Applied and Environmental Microbiology, 2002. **68**: p. 952-956.
110. Scholtz, J.M., et al., *Parameters of helix-coil transition theory for alanine-based peptides of varying chain lengths in water.* Biopolymers, 1991. **31**(13): p. 1463-70.
111. Hauge, H.H., et al., *Plantaracin A Is an Amphiphilic α-Helical Bacteriocin-like Pheromone Which Exerts Antimicrobial and Pheromone Activities through Different Mechanisms.* Biochemistry, 1998. **37**: p. 16026-16032.
112. Braunschweiler, L. and R.R. Ernst, J.Mag.Res., 1983. **53**: p. 521.
113. Jeener, J., et al., *Investigation of Exchange Processes by 2-Dimensional Nmr-Spectroscopy.* J.Chem.Phys., 1979. **71**(11): p. 4546-4553.
114. Wuthrich, K., *NMR of Proteins and Nucleic Acids*. 1986, New York: Wiley Interscience.
115. Sklenar, V., et al., *Gradient-Tailored Water Suppression for H-1-N-15 Hsqc Experiments Optimized to Retain Full Sensitivity.* J.Mag.Res. Ser. A, 1993. **102**(2): p. 241-245.
116. Delaglio, F., et al., *Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes.* J.Biomol. NMR, 1995. **6**(3): p. 277-293.
117. 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.
118. Koradi, R., M. Billeter, and K. Wuthrich, *MOLMOL: A program for display and analysis of macromolecular structures.* Journal of Molecular Graphics, 1996. **14**(1): p. 51.
119. Tichaczek, P.S., et al., *Characterization of the Bacteriocins Curvacin A and Sakacin P Produced by Lactobacillus curvatus LTH 1174 and L. sake LTH673.* Systematic and Applied Microbiology, 1992. **15**: p. 460-468.
120. Huth, J.R., et al., *Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR.* Protein Sci, 1997. **6**(11): p. 2359-64.
121. Koenig, B.W., M. Rogowski, and J.M. Louis, *A rapid method to attain isotope labelled small soluble peptides for NMR studies.* J Biomol NMR, 2003. **26**(3): p. 193-202.
122. Yan, L.Z., et al., *Analogues of bacteriocins: antimicrobial specificity and interactions of leucocin A with its enantiomer, carnobacteriocin B2, and truncated derivatives.* J Med Chem, 2000. **43**(24): p. 4579-81.
123. 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.
124. Marrink, S.J., D.P. Tieleman, and A.E. Mark, *Molecular dynamics simulations of micellization kinetics.* 2000.
125. Tieleman, D.P., S.J. Marrink, and H.J. Berendsen, *A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems.* Biochim Biophys Acta, 1997. **1331**(3): p. 235-70.
126. Wishart, D.S., B.D. Sykes, and F.M. Richards, *Simple Techniques for the Quantification of Protein Secondary Structure by H-1-Nmr Spectroscopy.* Febs Letters, 1991. **293**(1-2): p. 72-80.
127. 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.

128. Güntert, P., C. Mumenthaler, and T. Herrmann, *DYANA User's Manual, verison 1.5*. 1998, Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule: Zürich. p. 198.
129. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Res, 2000. **28**(1): p. 235-42.
130. Morisset, D., et al., *Mutational analysis of mesentericin y105, an anti-Listeria bacteriocin, for determination of impact on bactericidal activity, in vitro secondary structure, and membrane interaction.* Appl Environ Microbiol, 2004. **70**(8): p. 4672-80.
131. Lodish, H., et al., *Molecular Cell Biology*. 5.edition ed. 2003: W.H.Freeman. 1152.
132. Watts, A., *Solid-state NMR approaches for studying the interaction of peptides and proteins with membranes.* Biochim Biophys Acta, 1998. **1376**(3): p. 297-318.
133. Bechinger, B., et al., *Orientations of helical peptides in membrane bilayers by solid state NMR spectroscopy.* Solid State Nucl Magn Reson, 1996. **7**(3): p. 185-91.
134. Bechinger, B., et al., *Orientations of amphipathic helical peptides in membrane bilayers determined by solid-state NMR spectroscopy.* J Biomol NMR, 1991. **1**(2): p. 167-73.
135. Tian, C., et al., *Initial structural and dynamic characterization of the M2 protein transmembrane and amphipathic helices in lipid bilayers.* Protein Sci, 2003. **12**(11): p. 2597-605.
136. Nishimura, K., et al., *The closed state of a H+ channel helical bundle combining precise orientational and distance restraints from solid state NMR.* Biochemistry, 2002. **41**(44): p. 13170-7.
137. Zhao, H. and P.K. Kinnunen, *Binding of the antimicrobial peptide temporin L to liposomes assessed by Trp fluorescence.* J Biol Chem, 2002. **277**(28): p. 25170-7.
138. Fimland, G., V.G. Eijsink, and J. Nissen-Meyer, *Comparative studies of immunity proteins of pediocin-like bacteriocins.* Microbiology, 2002. **148**(Pt 11): p. 3661-70.

# 7 Appendix

## 7.1 Equipments and chemicals

| Equipment: | Manufacturer: |
|---|---|
| Wilmand NMR tubes (5mm) | Sigma-Aldrich |
| UV-visible recording spectrophotometer (UV-160) | Shimadzu |
| FPLC-system | Amersham Biosciences |
| Resource RPC (3mL) | Amersham Biosciences |
| Smart-system | Amersham Biosciences |
| Avanti Centrifuge J-25 | Beckman |
| JA-14 Beckman rotor | Beckman |
| Äkta purifier | Amersham Biosciences |
| Voyager-DE RP mass spectrometer | Perseptive biosystems |
| Biofuge pico | Heraeus instruments |
| J-810 Spectropolarimeter | Jasco |
| PCR-cycler | MJ Research |
| Histag binding column | Novagene |
| INOVA 800 NMR Spectrometer | Varian |
| 5 mm 1H {13C, 15N} pfg probe | Varian |

| Chemicals: | |
|---|---|
| Sodium 2,2-dimetyl-2-silapentan-5-sulfonat (DSS) | Cambridge Isotope Lab. |
| Deuterated dodecylphosphocholine (DPC) D38,9% | Cambridge Isotope Lab. |
| Deuterium oxide 99,96% | Aldrich Chemical Company |
| MRS Broth (CM 359) | Oxoid Ltd. |
| Propanol | Merck |
| NaCl | Merck |
| $NaH_2PO_4$ | Merck |
| $Na_2HPO_4$ | Merck |
| Triflourethanol (TFE) | Aldrich/Merck |
| Acetonitril | Merck |
| α-cyano-4-hydroxycinnamic acid (CHCA) | Sigma |
| Ethanol | Arcus |
| Q Sepharose Fast Flow | Amersham Biosciences |
| *E. coil* BL21(DE3)pLys S | Stratagene |
| pET 3a-plasmid | Novagene |
| Oglionucleotides | Eurogenetec |
| Restriction enzymes | Fermentas |
| Plasmid mini prep kit | QiAgene |
| GFX™ DNA purification kit | Amersham Biosciences |
| Mutagenesis kit | Stratagene |
| IPTG | Fermentas |
| Urea | Merck |
| TrisCl | Merck |
| DPC | CDN Isotopes |
| $D_2O$ | Cambridge Isotope Laboratories |

## 7.2 Resonance assignment of curvacin A

| RESIDUE 1 | | | RESIDUE 2 | | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 2 | ARG+ | HA | 2 | ARG+ | QG | 4.53 |
| 2 | ARG+ | HA | 2 | ARG+ | QB | 3.93 |
| 2 | ARG+ | HA | 2 | ARG+ | QD | 5.5 |
| 2 | ARG+ | HN | 2 | ARG+ | QB | 3.75 |
| 2 | ARG+ | HN | 2 | ARG+ | QG | 3.91 |
| 2 | ARG+ | HN | 2 | ARG+ | HA | 4.52 |
| 2 | ARG+ | QB | 2 | ARG+ | HE | 4.83 |
| 2 | ARG+ | QB | 2 | ARG+ | QD | 3.75 |
| 2 | ARG+ | QB | 2 | ARG+ | QG | 3.35 |
| 2 | ARG+ | QD | 2 | ARG+ | HE | 3.94 |
| 2 | ARG+ | QG | 2 | ARG+ | HE | 4.5 |
| 2 | ARG+ | QG | 2 | ARG+ | QD | 3.77 |
| 2 | ARG+ | HA | 3 | SER | HN | 3.64 |
| 2 | ARG+ | HN | 3 | SER | HN | 4.77 |
| 2 | ARG+ | QB | 3 | SER | HN | 4.9 |
| 2 | ARG+ | QG | 3 | SER | HN | 4.49 |
| 2 | ARG+ | QB | 4 | TYR | QE | 4.14 |
| 2 | ARG+ | QB | 4 | TYR | QD | 5.24 |
| 2 | ARG+ | QD | 4 | TYR | QE | 4.22 |
| 2 | ARG+ | QG | 4 | TYR | QE | 3.2 |
| 2 | ARG+ | QG | 4 | TYR | HB2 | 5.5 |
| 2 | ARG+ | QG | 9 | TYR | QD | 5.8 |
| 2 | ARG+ | QG | 9 | TYR | HB3 | 4.98 |
| 2 | ARG+ | QG | 9 | TYR | HA | 4.59 |
| 3 | SER | HA | 3 | SER | HB3 | 4.66 |
| 3 | SER | HA | 3 | SER | HB2 | 4.91 |
| 3 | SER | HB2 | 3 | SER | HB3 | 3.19 |
| 3 | SER | HN | 3 | SER | HB3 | 4.1 |
| 3 | SER | HN | 3 | SER | HB2 | 3.96 |
| 3 | SER | HN | 3 | SER | HA | 4.6 |
| 3 | SER | HA | 4 | TYR | HN | 3.6 |
| 3 | SER | HB2 | 4 | TYR | HN | 4.43 |
| 3 | SER | HB3 | 4 | TYR | HN | 4.43 |
| 3 | SER | HN | 4 | TYR | QD | 4.71 |
| 3 | SER | HN | 4 | TYR | HN | 3 |
| 3 | SER | HN | 8 | VAL | QG1 | 4.57 |
| 3 | SER | HN | 8 | VAL | QG2 | 4.19 |
| 3 | SER | HA | 9 | TYR | QD | 5.11 |
| 3 | SER | HA | 9 | TYR | QE | 5.5 |
| 3 | SER | HB2 | 9 | TYR | QE | 4.92 |
| 3 | SER | HB2 | 9 | TYR | QD | 5.5 |
| 3 | SER | HB3 | 9 | TYR | QE | 5.22 |
| 3 | SER | HN | 9 | TYR | QD | 5.28 |
| 4 | TYR | HB2 | 3 | SER | HN | 4 |

| | RESIDUE 1 | | | RESIDUE 2 | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 4 | TYR | HA | 4 | TYR | QD | 5.16 |
| 4 | TYR | HB2 | 4 | TYR | QD | 3.76 |
| 4 | TYR | HN | 4 | TYR | HB2 | 3.88 |
| 4 | TYR | HN | 4 | TYR | HA | 4.28 |
| 4 | TYR | QD | 4 | TYR | QE | 2.98 |
| 4 | TYR | HB2 | 8 | VAL | QG2 | 4.81 |
| 4 | TYR | HB3 | 8 | VAL | QG2 | 4.82 |
| 4 | TYR | HB3 | 8 | VAL | QG1 | 5.24 |
| 4 | TYR | HB3 | 8 | VAL | HA | 6 |
| 4 | TYR | QD | 8 | VAL | QG2 | 4.8 |
| 4 | TYR | QD | 8 | VAL | HB | 6 |
| 5 | GLY | HN | 5 | GLY | QA | 3.97 |
| 6 | ASN | HA | 6 | ASN | QB | 4.91 |
| 6 | ASN | HD21 | 6 | ASN | HD22 | 2.69 |
| 6 | ASN | HN | 6 | ASN | QB | 3.42 |
| 6 | ASN | QB | 6 | ASN | HD22 | 3.7 |
| 6 | ASN | QB | 6 | ASN | HD21 | 3.87 |
| 6 | ASN | QB | 7 | GLY | HN | 4.52 |
| 7 | GLY | HA2 | 7 | GLY | HA1 | 3.89 |
| 7 | GLY | HN | 7 | GLY | HA2 | 3.5 |
| 7 | GLY | HN | 7 | GLY | HA1 | 3.6 |
| 7 | GLY | HA1 | 8 | VAL | HN | 4.19 |
| 7 | GLY | HA2 | 8 | VAL | HN | 4.28 |
| 7 | GLY | HA2 | 8 | VAL | HB | 5.5 |
| 7 | GLY | HN | 8 | VAL | HN | 4.1 |
| 7 | GLY | HN | 8 | VAL | HB | 4.8 |
| 7 | GLY | HA1 | 9 | TYR | QE | 5.41 |
| 7 | GLY | HA2 | 9 | TYR | QE | 5.58 |
| 8 | VAL | QG1 | 4 | TYR | HN | 5 |
| 8 | VAL | HA | 8 | VAL | QG1 | 3.89 |
| 8 | VAL | HA | 8 | VAL | QG2 | 3.69 |
| 8 | VAL | HA | 8 | VAL | HB | 3.89 |
| 8 | VAL | HB | 8 | VAL | QG1 | 3.93 |
| 8 | VAL | HB | 8 | VAL | QG2 | 3.58 |
| 8 | VAL | HN | 8 | VAL | QG1 | 3.71 |
| 8 | VAL | HN | 8 | VAL | HB | 3.82 |
| 8 | VAL | HN | 8 | VAL | QG2 | 4.3 |
| 8 | VAL | HN | 8 | VAL | HA | 4.28 |
| 8 | VAL | HB | 9 | TYR | HN | 3.66 |
| 8 | VAL | HN | 9 | TYR | HN | 4.31 |
| 8 | VAL | QG1 | 9 | TYR | HN | 4.67 |
| 8 | VAL | QG2 | 9 | TYR | HN | 4.25 |
| 9 | TYR | QE | 7 | GLY | HN | 4.8 |
| 9 | TYR | QE | 8 | VAL | QG2 | 5.9 |
| 9 | TYR | HA | 9 | TYR | QD | 4.35 |
| 9 | TYR | HA | 9 | TYR | HB3 | 4.9 |
| 9 | TYR | HN | 9 | TYR | QE | 4.53 |
| 9 | TYR | HN | 9 | TYR | QD | 4.21 |

| RESIDUE 1 | | | RESIDUE 2 | | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 9 | TYR | HN | 9 | TYR | HB3 | 3.51 |
| 9 | TYR | HN | 9 | TYR | HA | 3.36 |
| 9 | TYR | QD | 9 | TYR | HB3 | 3.87 |
| 9 | TYR | QE | 9 | TYR | QD | 2.96 |
| 9 | TYR | QE | 9 | TYR | HB3 | 4.51 |
| 9 | TYR | QE | 9 | TYR | HB2 | 4.53 |
| 9 | TYR | HN | 10 | CYSS | HN | 3.62 |
| 9 | TYR | QD | 10 | CYSS | HN | 4.89 |
| 9 | TYR | QD | 10 | CYSS | HB3 | 6 |
| 9 | TYR | QD | 16 | TRP | HE1 | 6 |
| 10 | CYSS | HN | 9 | TYR | HB3 | 6 |
| 10 | CYSS | HN | 10 | CYSS | HB2 | 4.21 |
| 10 | CYSS | HN | 10 | CYSS | HB3 | 4.5 |
| 10 | CYSS | HB3 | 11 | ASN | HN | 5.5 |
| 10 | CYSS | HN | 11 | ASN | HN | 4.72 |
| 11 | ASN | HA | 11 | ASN | HB2 | 3.85 |
| 11 | ASN | HA | 11 | ASN | HB3 | 4.2 |
| 11 | ASN | HB2 | 11 | ASN | HD21 | 4.47 |
| 11 | ASN | HB2 | 11 | ASN | HD22 | 4.27 |
| 11 | ASN | HB2 | 11 | ASN | HB3 | 2.81 |
| 11 | ASN | HB3 | 11 | ASN | HD21 | 4.8 |
| 11 | ASN | HB3 | 11 | ASN | HD22 | 3.61 |
| 11 | ASN | HD21 | 11 | ASN | HD22 | 3 |
| 11 | ASN | HN | 11 | ASN | HB2 | 3.95 |
| 11 | ASN | HN | 11 | ASN | HB3 | 3.89 |
| 11 | ASN | HN | 11 | ASN | HA | 3.96 |
| 11 | ASN | HB2 | 12 | ASN | HD22 | 5.5 |
| 11 | ASN | HB2 | 12 | ASN | HN | 4.29 |
| 11 | ASN | HB3 | 12 | ASN | HN | 3.7 |
| 11 | ASN | HD21 | 12 | ASN | HD22 | 4.55 |
| 11 | ASN | HD21 | 12 | ASN | HB2 | 5.2 |
| 11 | ASN | HB2 | 14 | LYS+ | QZ | 5 |
| 11 | ASN | HB3 | 14 | LYS+ | QZ | 5.57 |
| 11 | ASN | HD22 | 14 | LYS+ | QE | 5 |
| 11 | ASN | HN | 14 | LYS+ | HN | 5 |
| 11 | ASN | HB2 | 16 | TRP | HH2 | 4.94 |
| 11 | ASN | HB2 | 16 | TRP | HZ3 | 5 |
| 11 | ASN | HB3 | 16 | TRP | HH2 | 5 |
| 11 | ASN | HB3 | 16 | TRP | HZ3 | 5 |
| 11 | ASN | HN | 16 | TRP | HZ3 | 5 |
| 12 | ASN | HB2 | 12 | ASN | HD21 | 3.95 |
| 12 | ASN | HB2 | 12 | ASN | HB3 | 2.72 |
| 12 | ASN | HB2 | 12 | ASN | HD22 | 4.14 |
| 12 | ASN | HB3 | 12 | ASN | HD21 | 4.12 |
| 12 | ASN | HB3 | 12 | ASN | HD22 | 4.29 |
| 12 | ASN | HD22 | 12 | ASN | HD21 | 2.79 |
| 12 | ASN | HN | 12 | ASN | HB3 | 4.7 |

| RESIDUE 1 | | | RESIDUE 2 | | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 12 | ASN | HN | 12 | ASN | HD21 | 4.23 |
| 12 | ASN | HN | 12 | ASN | HD22 | 5.38 |
| 12 | ASN | HN | 12 | ASN | HB2 | 3.78 |
| 12 | ASN | HN | 12 | ASN | HB3 | 4 |
| 12 | ASN | HD21 | 13 | LYS+ | QZ | 6 |
| 12 | ASN | HB3 | 14 | LYS+ | HN | 5.5 |
| 12 | ASN | HD21 | 14 | LYS+ | QE | 5.5 |
| 12 | ASN | HN | 14 | LYS+ | QZ | 5 |
| 13 | LYS+ | HA | 13 | LYS+ | QB | 3.31 |
| 13 | LYS+ | HA | 13 | LYS+ | HG2 | 3.9 |
| 13 | LYS+ | HA | 13 | LYS+ | QE | 5.43 |
| 13 | LYS+ | HA | 13 | LYS+ | QD | 6 |
| 13 | LYS+ | HG2 | 13 | LYS+ | QE | 4.5 |
| 13 | LYS+ | HG2 | 13 | LYS+ | QD | 3.24 |
| 13 | LYS+ | HG3 | 13 | LYS+ | QE | 4.91 |
| 13 | LYS+ | HG3 | 13 | LYS+ | QD | 3.27 |
| 13 | LYS+ | HN | 13 | LYS+ | HG2 | 4.73 |
| 13 | LYS+ | HN | 13 | LYS+ | HG3 | 5.24 |
| 13 | LYS+ | HN | 13 | LYS+ | HA | 3.74 |
| 13 | LYS+ | QB | 13 | LYS+ | QD | 3.9 |
| 13 | LYS+ | QB | 13 | LYS+ | HN | 4 |
| 13 | LYS+ | QB | 13 | LYS+ | HG2 | 4 |
| 13 | LYS+ | QD | 13 | LYS+ | QE | 3.53 |
| 13 | LYS+ | QD | 13 | LYS+ | QZ | 5.1 |
| 13 | LYS+ | QE | 13 | LYS+ | QZ | 3.93 |
| 13 | LYS+ | QE | 13 | LYS+ | QB | 6 |
| 13 | LYS+ | HA | 14 | LYS+ | QG | 5.7 |
| 13 | LYS+ | HG2 | 14 | LYS+ | HN | 4.36 |
| 13 | LYS+ | HN | 14 | LYS+ | HN | 4.29 |
| 13 | LYS+ | HN | 14 | LYS+ | QB | 5.5 |
| 13 | LYS+ | HN | 14 | LYS+ | QG | 5.5 |
| 13 | LYS+ | QB | 14 | LYS+ | QZ | 5.4 |
| 13 | LYS+ | QB | 14 | LYS+ | HN | 3.87 |
| 13 | LYS+ | QD | 14 | LYS+ | QE | 5.2 |
| 13 | LYS+ | QD | 14 | LYS+ | HN | 3.82 |
| 13 | LYS+ | QD | 14 | LYS+ | HN | 6 |
| 13 | LYS+ | QE | 14 | LYS+ | QB | 5.36 |
| 13 | LYS+ | QE | 14 | LYS+ | QD | 4.1 |
| 13 | LYS+ | QE | 14 | LYS+ | QG | 4.9 |
| 14 | LYS+ | QE | 13 | LYS+ | QB | 6 |
| 14 | LYS+ | QG | 13 | LYS+ | QB | 4 |
| 14 | LYS+ | QZ | 13 | LYS+ | QB | 6 |
| 14 | LYS+ | HA | 14 | LYS+ | QB | 3.69 |
| 14 | LYS+ | HA | 14 | LYS+ | QG | 4.55 |
| 14 | LYS+ | HA | 14 | LYS+ | QD | 5.27 |
| 14 | LYS+ | HN | 14 | LYS+ | QB | 3.61 |
| 14 | LYS+ | HN | 14 | LYS+ | QG | 3.95 |

| RESIDUE 1 | | | RESIDUE 2 | | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 14 | LYS+ | HN | 14 | LYS+ | HA | 4.41 |
| 14 | LYS+ | HN | 14 | LYS+ | QE | 6 |
| 14 | LYS+ | QB | 14 | LYS+ | QE | 4.1 |
| 14 | LYS+ | QB | 14 | LYS+ | QD | 3.29 |
| 14 | LYS+ | QB | 14 | LYS+ | QG | 3.34 |
| 14 | LYS+ | QB | 14 | LYS+ | QZ | 5.5 |
| 14 | LYS+ | QD | 14 | LYS+ | QE | 3.7 |
| 14 | LYS+ | QD | 14 | LYS+ | QZ | 5.5 |
| 14 | LYS+ | QE | 14 | LYS+ | QZ | 4.76 |
| 14 | LYS+ | QG | 14 | LYS+ | QZ | 4.43 |
| 14 | LYS+ | QG | 14 | LYS+ | QD | 3.16 |
| 14 | LYS+ | QG | 14 | LYS+ | QE | 4.63 |
| 14 | LYS+ | HA | 15 | CYSS | HN | 4.7 |
| 14 | LYS+ | QB | 15 | CYSS | HB2 | 5.5 |
| 14 | LYS+ | QB | 15 | CYSS | HB3 | 5.5 |
| 14 | LYS+ | QB | 15 | CYSS | HN | 3.83 |
| 14 | LYS+ | QE | 15 | CYSS | HN | 4.27 |
| 14 | LYS+ | QG | 15 | CYSS | HN | 4.33 |
| 14 | LYS+ | QZ | 15 | CYSS | HN | 5.15 |
| 14 | LYS+ | QB | 16 | TRP | HN | 5.2 |
| 15 | CYSS | HN | 11 | ASN | HB2 | 6 |
| 15 | CYSS | HA | 15 | CYSS | HB3 | 5.12 |
| 15 | CYSS | HN | 15 | CYSS | HB2 | 3.94 |
| 15 | CYSS | HN | 15 | CYSS | HB3 | 3.97 |
| 16 | TRP | HA | 16 | TRP | QB | 4.44 |
| 16 | TRP | HA | 16 | TRP | HE3 | 5.5 |
| 16 | TRP | HA | 16 | TRP | HD1 | 5 |
| 16 | TRP | HD1 | 16 | TRP | HE1 | 3.68 |
| 16 | TRP | HE1 | 16 | TRP | HH2 | 5.47 |
| 16 | TRP | HE3 | 16 | TRP | HH2 | 4.47 |
| 16 | TRP | HE3 | 16 | TRP | HZ3 | 3.46 |
| 16 | TRP | HE3 | 16 | TRP | HD1 | 5.2 |
| 16 | TRP | HE3 | 16 | TRP | HN | 5.8 |
| 16 | TRP | HH2 | 16 | TRP | HZ2 | 3.42 |
| 16 | TRP | HH2 | 16 | TRP | HZ3 | 3.4 |
| 16 | TRP | HN | 16 | TRP | HD1 | 5.5 |
| 16 | TRP | HN | 16 | TRP | QB | 3.72 |
| 16 | TRP | HN | 16 | TRP | HA | 3.72 |
| 16 | TRP | HZ2 | 16 | TRP | HD1 | 5.6 |
| 16 | TRP | HZ2 | 16 | TRP | HE1 | 4.21 |
| 16 | TRP | HZ3 | 16 | TRP | HZ2 | 4.32 |
| 16 | TRP | QB | 16 | TRP | HE3 | 3.68 |
| 16 | TRP | QB | 16 | TRP | HD1 | 3.82 |
| 16 | TRP | QB | 16 | TRP | HZ3 | 5.23 |
| 16 | TRP | HA | 17 | VAL | QG1 | 6 |
| 16 | TRP | HA | 17 | VAL | QG2 | 6 |
| 16 | TRP | HD1 | 17 | VAL | QG1 | 5.5 |

| RESIDUE 1 | | | RESIDUE 2 | | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 16 | TRP | HN | 17 | VAL | QG1 | 4.64 |
| 16 | TRP | QB | 17 | VAL | QG1 | 4.87 |
| 16 | TRP | QB | 17 | VAL | QG2 | 6 |
| 16 | TRP | HD1 | 19 | ARG+ | HB2 | 6 |
| 17 | VAL | HA | 17 | VAL | QG1 | 3.41 |
| 17 | VAL | HA | 17 | VAL | HB | 3.66 |
| 17 | VAL | HA | 17 | VAL | QG2 | 4 |
| 17 | VAL | HN | 17 | VAL | QG1 | 5 |
| 17 | VAL | QG1 | 17 | VAL | HB | 2.98 |
| 17 | VAL | HB | 18 | ASN | HN | 5.5 |
| 17 | VAL | QG1 | 18 | ASN | HN | 4.24 |
| 17 | VAL | QG1 | 18 | ASN | HB2 | 4.83 |
| 17 | VAL | QG2 | 18 | ASN | HB2 | 5.2 |
| 18 | ASN | HB2 | 18 | ASN | HD21 | 3.77 |
| 18 | ASN | HB3 | 18 | ASN | HD22 | 5.5 |
| 18 | ASN | HB3 | 18 | ASN | HD21 | 4.5 |
| 18 | ASN | HB3 | 18 | ASN | HB2 | 2.94 |
| 18 | ASN | HD22 | 18 | ASN | HD21 | 2.82 |
| 18 | ASN | HN | 18 | ASN | HB3 | 4.19 |
| 18 | ASN | HN | 18 | ASN | HB2 | 3.73 |
| 18 | ASN | HN | 18 | ASN | HB3 | 4 |
| 19 | ARG+ | QD | 6 | ASN | HD22 | 5 |
| 19 | ARG+ | QH1 | 6 | ASN | HD22 | 5 |
| 19 | ARG+ | QD | 16 | TRP | HZ3 | 6 |
| 19 | ARG+ | HA | 19 | ARG+ | HB2 | 3.36 |
| 19 | ARG+ | HA | 19 | ARG+ | QG | 3.98 |
| 19 | ARG+ | HA | 19 | ARG+ | HB3 | 5.6 |
| 19 | ARG+ | HA | 19 | ARG+ | QD | 5.26 |
| 19 | ARG+ | HA | 19 | ARG+ | HE | 4.96 |
| 19 | ARG+ | HB2 | 19 | ARG+ | HE | 4.24 |
| 19 | ARG+ | HB2 | 19 | ARG+ | QG | 3.18 |
| 19 | ARG+ | HB3 | 19 | ARG+ | HE | 4.23 |
| 19 | ARG+ | HB3 | 19 | ARG+ | QD | 4.32 |
| 19 | ARG+ | HN | 19 | ARG+ | QD | 5.5 |
| 19 | ARG+ | HN | 19 | ARG+ | HA | 3.91 |
| 19 | ARG+ | HN | 19 | ARG+ | HB3 | 4.9 |
| 19 | ARG+ | HN | 19 | ARG+ | HB2 | 3.52 |
| 19 | ARG+ | HN | 19 | ARG+ | QG | 4.1 |
| 19 | ARG+ | QD | 19 | ARG+ | HE | 3.9 |
| 19 | ARG+ | QD | 19 | ARG+ | QG | 4.4 |
| 19 | ARG+ | QD | 19 | ARG+ | HB2 | 4.71 |
| 19 | ARG+ | QD | 19 | ARG+ | HB2 | 3.97 |
| 19 | ARG+ | QD | 19 | ARG+ | HB3 | 4.1 |
| 19 | ARG+ | QG | 19 | ARG+ | HE | 4.2 |
| 19 | ARG+ | QG | 19 | ARG+ | QD | 4.62 |
| 19 | ARG+ | HA | 20 | GLY | HN | 4.36 |
| 19 | ARG+ | HB2 | 20 | GLY | QA | 5.5 |

| RESIDUE 1 | | | RESIDUE 2 | | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 19 | ARG+ | HN | 20 | GLY | HN | 4.33 |
| 19 | ARG+ | QG | 20 | GLY | HN | 5.24 |
| 19 | ARG+ | HA | 22 | ALA | QB | 6 |
| 20 | GLY | HN | 20 | GLY | QA | 3.4 |
| 20 | GLY | HN | 21 | GLU | HN | 4.36 |
| 20 | GLY | QA | 21 | GLU | HN | 3.72 |
| 21 | GLU | HA | 21 | GLU | HG2 | 4.25 |
| 21 | GLU | HA | 21 | GLU | HG3 | 4.2 |
| 21 | GLU | HA | 21 | GLU | HB2 | 4.31 |
| 21 | GLU | HA | 21 | GLU | HB3 | 4.38 |
| 21 | GLU | HB2 | 21 | GLU | HG3 | 4.33 |
| 21 | GLU | HB2 | 21 | GLU | HG2 | 4.49 |
| 21 | GLU | HB3 | 21 | GLU | HG2 | 3.99 |
| 21 | GLU | HB3 | 21 | GLU | HG3 | 4.33 |
| 21 | GLU | HG2 | 21 | GLU | HG3 | 3.35 |
| 21 | GLU | HN | 21 | GLU | HB3 | 3.58 |
| 21 | GLU | HN | 21 | GLU | HB2 | 3.54 |
| 21 | GLU | HN | 21 | GLU | HG2 | 6 |
| 21 | GLU | HN | 21 | GLU | HG3 | 6 |
| 21 | GLU | HB3 | 22 | ALA | HN | 4.27 |
| 21 | GLU | HG3 | 22 | ALA | HN | 5.5 |
| 21 | GLU | HG3 | 22 | ALA | QB | 5.5 |
| 21 | GLU | HA | 24 | GLN | HN | 5.5 |
| 22 | ALA | HA | 22 | ALA | QB | 3.42 |
| 22 | ALA | HN | 22 | ALA | HA | 3.91 |
| 22 | ALA | HN | 22 | ALA | QB | 3.41 |
| 22 | ALA | HN | 23 | THR | HN | 3.72 |
| 22 | ALA | QB | 23 | THR | HB | 4.18 |
| 22 | ALA | QB | 23 | THR | HN | 3.77 |
| 22 | ALA | QB | 24 | GLN | QG | 5.42 |
| 23 | THR | HA | 23 | THR | QG2 | 3.49 |
| 23 | THR | HA | 23 | THR | HB | 4.99 |
| 23 | THR | HN | 23 | THR | HA | 4.9 |
| 23 | THR | HN | 23 | THR | HB | 3.65 |
| 23 | THR | HA | 24 | GLN | HN | 4.54 |
| 23 | THR | HN | 24 | GLN | HN | 3.71 |
| 23 | THR | HN | 24 | GLN | QG | 4.2 |
| 23 | THR | QG2 | 24 | GLN | HN | 5.18 |
| 24 | GLN | HN | 22 | ALA | QB | 5 |
| 24 | GLN | HA | 24 | GLN | QG | 3.88 |
| 24 | GLN | HA | 24 | GLN | HB3 | 3.86 |
| 24 | GLN | HA | 24 | GLN | HB2 | 3.95 |
| 24 | GLN | HB2 | 24 | GLN | QG | 3.75 |
| 24 | GLN | HB2 | 24 | GLN | HB3 | 3.25 |
| 24 | GLN | HB3 | 24 | GLN | QG | 3.83 |
| 24 | GLN | HE22 | 24 | GLN | HE21 | 3.12 |
| 24 | GLN | HN | 24 | GLN | HB2 | 4 |

| | RESIDUE 1 | | | RESIDUE 2 | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 24 | GLN | HN | 24 | GLN | HA | 3.61 |
| 24 | GLN | HN | 24 | GLN | HB3 | 3.65 |
| 24 | GLN | HN | 24 | GLN | QG | 4.5 |
| 24 | GLN | QG | 24 | GLN | HE21 | 4.32 |
| 24 | GLN | QG | 24 | GLN | HE22 | 4.62 |
| 24 | GLN | HB2 | 25 | SER | HN | 6 |
| 24 | GLN | HB3 | 25 | SER | HN | 6 |
| 25 | SER | HA | 25 | SER | HB3 | 4.45 |
| 25 | SER | HA | 25 | SER | HB2 | 5 |
| 25 | SER | HB2 | 25 | SER | HB3 | 3.3 |
| 25 | SER | HN | 25 | SER | HB3 | 3.69 |
| 25 | SER | HN | 25 | SER | HA | 4.14 |
| 25 | SER | HB2 | 26 | ILE | QG1 | 6 |
| 25 | SER | HB3 | 26 | ILE | HN | 4.2 |
| 25 | SER | HB3 | 27 | ILE | HN | 3.98 |
| 26 | ILE | QG2 | 25 | SER | HB3 | 5.58 |
| 26 | ILE | HA | 26 | ILE | QD1 | 4 |
| 26 | ILE | HA | 26 | ILE | HB | 3.78 |
| 26 | ILE | HA | 26 | ILE | QG2 | 5.32 |
| 26 | ILE | HB | 26 | ILE | QG1 | 3.87 |
| 26 | ILE | HB | 26 | ILE | QD1 | 3.92 |
| 26 | ILE | HN | 26 | ILE | QG2 | 4.37 |
| 26 | ILE | HN | 26 | ILE | QD1 | 4.42 |
| 26 | ILE | HN | 26 | ILE | HA | 3.69 |
| 26 | ILE | HN | 26 | ILE | QG1 | 4.17 |
| 26 | ILE | HN | 26 | ILE | HB | 3.61 |
| 26 | ILE | HA | 27 | ILE | HN | 5.5 |
| 26 | ILE | HB | 27 | ILE | HN | 3.86 |
| 26 | ILE | QG1 | 27 | ILE | HN | 6 |
| 26 | ILE | QG2 | 27 | ILE | HN | 4.24 |
| 26 | ILE | QG2 | 27 | ILE | HA | 5.46 |
| 26 | ILE | QG2 | 27 | ILE | QG2 | 6 |
| 27 | ILE | HA | 27 | ILE | QG1 | 4.35 |
| 27 | ILE | HA | 27 | ILE | HB | 4.4 |
| 27 | ILE | HA | 27 | ILE | QG2 | 3.6 |
| 27 | ILE | HB | 27 | ILE | QD1 | 3.65 |
| 27 | ILE | HB | 27 | ILE | QG1 | 3.6 |
| 27 | ILE | HN | 27 | ILE | HB | 3.43 |
| 27 | ILE | HN | 27 | ILE | HA | 3.63 |
| 27 | ILE | HN | 27 | ILE | QD1 | 3.87 |
| 27 | ILE | HN | 27 | ILE | QG2 | 3.66 |
| 27 | ILE | HN | 27 | ILE | QG1 | 3.9 |
| 27 | ILE | QG2 | 27 | ILE | QG1 | 3.14 |
| 27 | ILE | QG2 | 27 | ILE | HB | 3.8 |
| 27 | ILE | HN | 28 | GLY | HN | 3.73 |
| 28 | GLY | HA1 | 28 | GLY | HA2 | 3.97 |
| 28 | GLY | HN | 28 | GLY | HA2 | 3.76 |

| RESIDUE 1 | | | RESIDUE 2 | | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 28 | GLY | HN | 28 | GLY | HA1 | 3.6 |
| 28 | GLY | HA2 | 29 | GLY | HN | 4.46 |
| 28 | GLY | HN | 29 | GLY | HN | 4.11 |
| 28 | GLY | HN | 30 | MET | HN | 6 |
| 29 | GLY | HA1 | 29 | GLY | HA2 | 3.28 |
| 29 | GLY | HN | 29 | GLY | HA2 | 4.8 |
| 29 | GLY | HN | 29 | GLY | HA1 | 3.19 |
| 29 | GLY | HN | 30 | MET | HN | 4.9 |
| 30 | MET | HB2 | 27 | ILE | HA | |
| 30 | MET | HA | 30 | MET | HG2 | 4.29 |
| 30 | MET | HA | 30 | MET | HG3 | 3.99 |
| 30 | MET | HA | 30 | MET | HB2 | 3.9 |
| 30 | MET | HA | 30 | MET | HB3 | 3.54 |
| 30 | MET | HB2 | 30 | MET | HG2 | 3.86 |
| 30 | MET | HB2 | 30 | MET | HG3 | 3.67 |
| 30 | MET | HB3 | 30 | MET | HG3 | 3.92 |
| 30 | MET | HB3 | 30 | MET | HB2 | 3.16 |
| 30 | MET | HB3 | 30 | MET | HG2 | 4.2 |
| 30 | MET | HG2 | 30 | MET | HG3 | 3.14 |
| 30 | MET | HN | 30 | MET | HA | 3.91 |
| 30 | MET | HN | 30 | MET | HB3 | 3.92 |
| 30 | MET | HN | 30 | MET | HB2 | 3.81 |
| 30 | MET | HN | 30 | MET | HG2 | 4.16 |
| 30 | MET | HN | 30 | MET | HG3 | 4.43 |
| 30 | MET | HA | 31 | ILE | HN | 4.22 |
| 30 | MET | HB2 | 31 | ILE | HA | 4.59 |
| 30 | MET | HB2 | 31 | ILE | HN | 4.36 |
| 30 | MET | HB3 | 31 | ILE | HN | 4.3 |
| 30 | MET | HG2 | 31 | ILE | HN | 5.12 |
| 30 | MET | HG3 | 31 | ILE | HN | 5.49 |
| 30 | MET | HN | 31 | ILE | HN | 3.56 |
| 31 | ILE | HA | 31 | ILE | QG1 | 3.34 |
| 31 | ILE | HA | 31 | ILE | HB | 3.84 |
| 31 | ILE | HA | 31 | ILE | QD1 | 4 |
| 31 | ILE | HB | 31 | ILE | QG1 | 3.2 |
| 31 | ILE | HN | 31 | ILE | QD1 | 4.51 |
| 31 | ILE | HN | 31 | ILE | HA | 4.21 |
| 31 | ILE | HN | 31 | ILE | HB | 4.9 |
| 31 | ILE | HN | 31 | ILE | QG2 | 3.84 |
| 31 | ILE | HN | 31 | ILE | QG1 | 4.2 |
| 31 | ILE | QG2 | 31 | ILE | QG1 | 3.14 |
| 31 | ILE | QG2 | 31 | ILE | HB | 4.14 |
| 31 | ILE | QG2 | 31 | ILE | QD1 | 2.65 |
| 31 | ILE | HA | 32 | SER | HN | 4.36 |
| 31 | ILE | HB | 32 | SER | HN | 4 |
| 31 | ILE | HN | 32 | SER | HN | 3.65 |
| 31 | ILE | QD1 | 32 | SER | HN | 4.16 |

| RESIDUE 1 | | | RESIDUE 2 | | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 31 | ILE | QG1 | 32 | SER | HN | 3.41 |
| 31 | ILE | QG1 | 32 | SER | HA | 3.91 |
| 31 | ILE | QG2 | 32 | SER | HB3 | 5.1 |
| 31 | ILE | QG2 | 32 | SER | HA | 3.58 |
| 31 | ILE | QG2 | 32 | SER | HN | 4.19 |
| 31 | ILE | HB | 33 | GLY | HN | 6 |
| 31 | ILE | HA | 34 | TRP | HB2 | 5.2 |
| 31 | ILE | HA | 34 | TRP | HE3 | 4.4 |
| 31 | ILE | HA | 34 | TRP | HB3 | 4.39 |
| 31 | ILE | HA | 34 | TRP | HN | 3.95 |
| 31 | ILE | HA | 34 | TRP | HD1 | 6 |
| 31 | ILE | QG1 | 34 | TRP | HE3 | 4.72 |
| 31 | ILE | QG1 | 34 | TRP | HN | 6 |
| 31 | ILE | QG2 | 34 | TRP | HE3 | 5.5 |
| 31 | ILE | HA | 35 | ALA | QB | 6 |
| 31 | ILE | QG1 | 35 | ALA | HN | 4.75 |
| 32 | SER | HA | 32 | SER | HB2 | 5.19 |
| 32 | SER | HA | 32 | SER | HB3 | 6 |
| 32 | SER | HA | 32 | SER | HB3 | 4 |
| 32 | SER | HN | 32 | SER | HA | 3.35 |
| 32 | SER | HN | 32 | SER | HB3 | 3.7 |
| 32 | SER | HN | 34 | TRP | HN | 5 |
| 32 | SER | HA | 35 | ALA | HN | 5.2 |
| 32 | SER | HA | 35 | ALA | QB | 5.5 |
| 32 | SER | HA | 36 | SER | HN | 5.5 |
| 33 | GLY | HN | 30 | MET | HB3 | 5.8 |
| 33 | GLY | HN | 33 | GLY | QA | 4.26 |
| 33 | GLY | HN | 34 | TRP | HN | 3.78 |
| 33 | GLY | HN | 35 | ALA | HN | 5.5 |
| 34 | TRP | HA | 34 | TRP | HD1 | 4.62 |
| 34 | TRP | HB2 | 34 | TRP | HD1 | 4.57 |
| 34 | TRP | HB2 | 34 | TRP | HE3 | 4.35 |
| 34 | TRP | HB3 | 34 | TRP | HD1 | 4.34 |
| 34 | TRP | HB3 | 34 | TRP | HE3 | 3.86 |
| 34 | TRP | HB3 | 34 | TRP | HZ3 | 5.5 |
| 34 | TRP | HE1 | 34 | TRP | HD1 | 3.82 |
| 34 | TRP | HE3 | 34 | TRP | HD1 | 5.16 |
| 34 | TRP | HE3 | 34 | TRP | HH2 | 5.2 |
| 34 | TRP | HE3 | 34 | TRP | HZ3 | 3.37 |
| 34 | TRP | HH2 | 34 | TRP | HZ3 | 3.3 |
| 34 | TRP | HN | 34 | TRP | HD1 | 4.67 |
| 34 | TRP | HN | 34 | TRP | HB3 | 3.83 |
| 34 | TRP | HN | 34 | TRP | HB2 | 3.75 |
| 34 | TRP | HN | 34 | TRP | HA | 4.41 |
| 34 | TRP | HN | 34 | TRP | HD1 | 6 |
| 34 | TRP | HN | 34 | TRP | HE3 | 6 |
| 34 | TRP | HZ2 | 34 | TRP | HD1 | 5.4 |

| RESIDUE 1 | | | RESIDUE 2 | | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 34 | TRP | HZ2 | 34 | TRP | HH2 | 3.44 |
| 34 | TRP | HZ2 | 34 | TRP | HE1 | 4.42 |
| 34 | TRP | HZ3 | 34 | TRP | HZ2 | 4.28 |
| 34 | TRP | HB2 | 35 | ALA | HN | 4.46 |
| 34 | TRP | HB2 | 35 | ALA | QB | 5.17 |
| 34 | TRP | HB3 | 35 | ALA | QB | 5.18 |
| 34 | TRP | HB3 | 35 | ALA | HN | 4.22 |
| 34 | TRP | HD1 | 35 | ALA | HN | 6 |
| 34 | TRP | HE3 | 35 | ALA | HA | 4.7 |
| 34 | TRP | HE3 | 35 | ALA | HN | 4.73 |
| 34 | TRP | HE3 | 35 | ALA | QB | 4.8 |
| 34 | TRP | HN | 35 | ALA | HN | 3.81 |
| 34 | TRP | HZ3 | 35 | ALA | QB | 5.35 |
| 34 | TRP | HB3 | 36 | SER | HN | 4.87 |
| 34 | TRP | HN | 36 | SER | HN | 3.84 |
| 34 | TRP | HE3 | 38 | LEU | QQD | 5.2 |
| 34 | TRP | HZ3 | 38 | LEU | QQD | 5.12 |
| 35 | ALA | HA | 35 | ALA | QB | 3.29 |
| 35 | ALA | HN | 35 | ALA | QB | 3.55 |
| 35 | ALA | HN | 35 | ALA | HA | 4.11 |
| 35 | ALA | HA | 36 | SER | HN | 3.62 |
| 35 | ALA | HN | 36 | SER | HN | 4.6 |
| 35 | ALA | QB | 36 | SER | HN | 3.84 |
| 35 | ALA | HA | 37 | GLY | HN | 4.53 |
| 35 | ALA | QB | 37 | GLY | HN | 4.5 |
| 35 | ALA | HA | 38 | LEU | HB2 | 5.5 |
| 35 | ALA | HA | 39 | ALA | HN | 5.5 |
| 36 | SER | HA | 36 | SER | QB | 4.4 |
| 36 | SER | HN | 36 | SER | HA | 3.9 |
| 36 | SER | HN | 36 | SER | QB | 3.56 |
| 36 | SER | HA | 37 | GLY | HN | 4.59 |
| 36 | SER | HN | 37 | GLY | HN | 3.45 |
| 36 | SER | HN | 38 | LEU | HN | 4.55 |
| 37 | GLY | HA1 | 37 | GLY | HA2 | 2.91 |
| 37 | GLY | HN | 37 | GLY | HA2 | 3.75 |
| 37 | GLY | HN | 37 | GLY | HA1 | 3.46 |
| 37 | GLY | HA1 | 38 | LEU | HN | 4.3 |
| 37 | GLY | HA2 | 38 | LEU | HN | 3.92 |
| 37 | GLY | HN | 38 | LEU | HG | 5.29 |
| 37 | GLY | HN | 38 | LEU | HN | 3.81 |
| 37 | GLY | HN | 38 | LEU | QQD | 4.36 |
| 37 | GLY | HA1 | 40 | GLY | HN | 5.5 |
| 37 | GLY | HA2 | 40 | GLY | HN | 5.5 |
| 38 | LEU | HB3 | 35 | ALA | HA | 5.5 |
| 38 | LEU | HA | 38 | LEU | QQD | 3.41 |
| 38 | LEU | HA | 38 | LEU | HB2 | 4.16 |
| 38 | LEU | HA | 38 | LEU | HG | 3.72 |

| RESIDUE 1 | | | RESIDUE 2 | | | DISTANCE |
|---|---|---|---|---|---|---|
| RESIDUENUMBER | RESIDUE | MOLECULE | RESIDUENUMBER | RESIDUE | MOLECULE | |
| 38 | LEU | HB2 | 38 | LEU | QQD | 3.39 |
| 38 | LEU | HB3 | 38 | LEU | HN | 3.61 |
| 38 | LEU | HG | 38 | LEU | QQD | 3.15 |
| 38 | LEU | HG | 38 | LEU | HB2 | 2.92 |
| 38 | LEU | HN | 38 | LEU | HB2 | 3.71 |
| 38 | LEU | HN | 38 | LEU | HA | 4.8 |
| 38 | LEU | HN | 38 | LEU | HG | 3.89 |
| 38 | LEU | HN | 38 | LEU | QQD | 3.88 |
| 38 | LEU | HG | 40 | GLY | HN | 5.5 |
| 38 | LEU | HN | 40 | GLY | HN | 5.5 |
| 39 | ALA | HA | 39 | ALA | QB | 3.84 |
| 39 | ALA | HN | 39 | ALA | QB | 3.49 |
| 39 | ALA | HN | 39 | ALA | HA | 3.93 |
| 39 | ALA | HA | 40 | GLY | HN | 4.26 |
| 39 | ALA | HN | 40 | GLY | HN | 4.42 |
| 39 | ALA | QB | 40 | GLY | HA1 | 3.87 |
| 39 | ALA | QB | 40 | GLY | HN | 3.91 |
| 39 | ALA | HA | 41 | MET | HN | 4.53 |
| 39 | ALA | QB | 41 | MET | HN | 4.49 |
| 40 | GLY | HN | 38 | LEU | QQD | 6 |
| 40 | GLY | HN | 40 | GLY | HA1 | 3.3 |
| 40 | GLY | HA1 | 41 | MET | HN | 4.54 |
| 40 | GLY | HA1 | 41 | MET | QB | 4.45 |
| 40 | GLY | HA2 | 41 | MET | HN | 4.16 |
| 40 | GLY | HN | 41 | MET | HN | 3.95 |
| 41 | MET | HA | 41 | MET | QB | 3.78 |
| 41 | MET | HA | 41 | MET | HG3 | 5.25 |
| 41 | MET | HG2 | 41 | MET | HG3 | 2.93 |
| 41 | MET | HN | 41 | MET | QB | 3.61 |
| 41 | MET | HN | 41 | MET | HA | 4.49 |
| 41 | MET | HN | 41 | MET | HG3 | 4.62 |
| 41 | MET | HN | 41 | MET | HG2 | 4.58 |
| 41 | MET | QB | 41 | MET | HG2 | 4.4 |