NMR as a Tool in Drug Research

Structure elucidation of peptidomimetics and pilicide-chaperone complexes

by

Mattias Hedenström

---

Akademisk avhandling

Som med tillstånd av rektorsämbetet vid Umeå universitet för erhållande av Filosofie Doktorsexamen vid Teknisk–Naturvetenskapliga fakulteten, framlägges till offentlig granskning vid Kemiska institutionen, Umeå universitet, sal KB3B1, KBC, fredagen den 8 oktober 2004, kl. 10.00.

Fakultetsopponent: Prof. Michael P. Williamson, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, England.
Title
NMR as a Tool in Drug Research:
Structure elucidation of peptidomimetics and pilicide-chaperone complexes

Author
Mattias Hedenström, Department of Chemistry, Organic Chemistry, Umeå University, SE–901 87, Umeå, Sweden.

Abstract
In the last decades NMR spectroscopy has become an invaluable tool both in academic research and in the pharmaceutical industry. This thesis describes applications of NMR spectroscopy in biomedical research for structure elucidation of biologically active peptides and peptidomimetics as well as in studies of ligand-protein interactions.

The first part of this thesis describes the theory and methodology of structure calculations of peptides using experimental restraints derived from NMR spectroscopy. This methodology has been applied to novel mimetics of the peptide hormones desmopressin and Leu-enkephalin. The results of these studies highlight the complicating issue of conformational exchange often encountered in structural determination of peptides and how careful analysis of experimental data as well as optimization of experimental conditions can enable structure determinations in such instances. Although the mimetics of both desmopressin and Leu-enkephalin were found to adopt the wanted conformations, they exhibited no or very poor biological activity. These results demonstrate the difficulties in designing peptidomimetics without detailed structural information of the receptors. A stereoselective synthetic route towards XxxΨ[CH₂O]Ala pseudodipeptides is also presented. Such pseudodipeptides can be used as isosteric amide bond replacements in peptides in order to increase their resistance towards proteolytic degradation.

The second part of this thesis describes the study of the interaction between compounds that inhibit pilus assembly, pilicides, and periplasmic chaperones from uropathogenic Escherichia coli. Periplasmic chaperones are key components in assembly of pili, i.e. hair-like protein complexes located on the surface of Escherichia coli that cause urinary tract infections. Detailed knowledge about this interaction is important in understanding how pilicides can inhibit pilus assembly by binding to chaperones. Relaxation-edited NMR experiments were used to confirm the affinity of the pilicides for the chaperones and chemical shift mapping was used to study the pilicide-chaperone interaction surface. These studies show that at least two interaction sites are present on the chaperone surface and consequently that two different mechanisms resulting in inhibition of pilus assembly may exist.

Keywords
NMR, structure calculations, peptidomimetics, desmopressin, Leu-enkephalin, chemical shift mapping, pilicides, chaperones, urinary tract infections.

Copyright © 2004 Mattias Hedenström
Printed in Sweden by Solfjäderm Offset AB, Umeå 2004
Contents

1. List of Papers ____________________________________________ 5
2. List of Abbreviations _____________________________________ 6
3. Nuclear Magnetic resonance _________________________________ 7
   3.1. A Brief History of Nuclear Magnetic Resonance _________ 7
   3.2. Basic NMR Theory _____________________________________ 8
   3.3. Nuclear Overhauser Effect ______________________________ 14
4. Structure Calculations of Peptides __________________________ 16
   4.1. Assignment of Peptides ________________________________ 17
   4.2. Distance and Angle Restraints _________________________ 18
   4.3. Other Experimental Restraints _________________________ 21
   4.4. Structure Calculations ________________________________ 22
5. Analogues of Biologically Active Peptides _________________ 25
   5.1. Turn Mimetics and Amidebond Isosteres _________________ 25
   5.2. Studies of Desmopressin ______________________________ 27
       5.2.1. Biological Background _____________________________ 27
       5.2.2. Stereoselective Synthesis of Ψ[CH2O] Pseudodipeptides 31
       5.2.3. Structure determination and biological evaluation of desmopressin analogues _____________________________ 33
   5.3. Studies of Leu-enkephalin _____________________________ 39
       5.3.1. Biological Background _____________________________ 39
       5.3.2. Design and synthesis of a 10-membered ring β-turn mimetic _________________________________ 40
       5.3.3. Structure determination and biological evaluation of a β-turn containing Leu-enkephalin analogue ____________ 42
   5.4. Conclusions _________________________________________ 44
6. Interaction between pilicides and chaperones involved in pilius assembly in uropathogenic E. coli __________________________ 45
   6.1. Biological Background _________________________________ 45
   6.2. NMR Studies of Chaperone-Pilicide Complexes __________ 49
       6.2.1. Relaxation-edited experiments ______________________ 49
       6.2.2. Assignment of FimC and Chemical Shift Mapping____ 51
1. **List of Papers**

This thesis is based on the papers listed below, which will be referred to in the text by the corresponding Roman numerals (I-V). Papers I-IV are reprinted with kind permission from the publishers.


V **M. Hedenström**, H. Emtenäs, N. Pemberton, S. Hultgren, J. Pinkner, V. Tegman, F. Almqvist, I. Sethson and J. Kihlberg, *NMR studies of the interactions between pilicides and periplasmic chaperones from uropathogenic E. coli*, Manuscript
# List of Abbreviations

## Abbreviation Meaning

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM</td>
<td>carboxymethyl</td>
</tr>
<tr>
<td>COSY</td>
<td>correlation spectroscopy</td>
</tr>
<tr>
<td>CW</td>
<td>continuous wave</td>
</tr>
<tr>
<td>DEAE</td>
<td>diethylaminoethyl</td>
</tr>
<tr>
<td>DG</td>
<td>distance geometry</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-dimethylaminopyridine</td>
</tr>
<tr>
<td>DMF</td>
<td>$N,N$-dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DQF</td>
<td>double quantum filter</td>
</tr>
<tr>
<td>FID</td>
<td>free induction decay</td>
</tr>
<tr>
<td>Fmoc</td>
<td>9-fluorenylmethoxycarbonyl</td>
</tr>
<tr>
<td>FT</td>
<td>fourier transform</td>
</tr>
<tr>
<td>Gal</td>
<td>galactose</td>
</tr>
<tr>
<td>GPCR</td>
<td>G-protein coupled receptor</td>
</tr>
<tr>
<td>HSQC</td>
<td>heteronuclear single quantum coherence spectroscopy</td>
</tr>
<tr>
<td>Hyl</td>
<td>hydroxylysine</td>
</tr>
<tr>
<td>Ig</td>
<td>immunoglobulin</td>
</tr>
<tr>
<td>IR</td>
<td>infrared</td>
</tr>
<tr>
<td>ISPA</td>
<td>isolated spin-pair approximation</td>
</tr>
<tr>
<td>MeOH</td>
<td>methanol</td>
</tr>
<tr>
<td>Mpa</td>
<td>mercaptopropionic acid</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>NOE</td>
<td>nuclear Overhauser effect</td>
</tr>
<tr>
<td>NOESY</td>
<td>nuclear Overhauser effect spectroscopy</td>
</tr>
<tr>
<td>PCA</td>
<td>principal component analysis</td>
</tr>
<tr>
<td>RF</td>
<td>radio frequency</td>
</tr>
<tr>
<td>rmsd</td>
<td>root mean square deviation</td>
</tr>
<tr>
<td>ROESY</td>
<td>rotating frame spectroscopy</td>
</tr>
<tr>
<td>SA</td>
<td>simulated annealing</td>
</tr>
<tr>
<td>STD</td>
<td>saturation transfer difference</td>
</tr>
<tr>
<td>$t$Bu</td>
<td>tert-butyl</td>
</tr>
<tr>
<td>TFE</td>
<td>trifluoroethanol</td>
</tr>
<tr>
<td>TOCSY</td>
<td>total correlation spectroscopy</td>
</tr>
<tr>
<td>UTI</td>
<td>urinary tract infection</td>
</tr>
</tbody>
</table>
3. Nuclear Magnetic resonance

3.1. A Brief History of Nuclear Magnetic Resonance

All spectroscopic techniques are by definition concerned with the interaction between atoms or molecules with electromagnetic radiation. More specifically, Nuclear Magnetic Resonance (NMR) spectroscopy is the study of how atomic nuclei absorb and emit electromagnetic radiation in presence of a static magnetic field. The theory of nuclear magnetic resonance dates back to the 1930’s, but at the time, all efforts to detect the phenomenon experimentally were fruitless.[1] In 1945, Ed Purcell and coworkers at Harvard University placed a piece of paraffin in a magnetic field for several hours and irradiated it with electromagnetic radiation of different frequencies. At a specific frequency, the protons in the paraffin sample emitted a detectable signal.[2] Only a few weeks later, Felix Bloch and coworkers at Stanford University were successful in their attempts to detect nuclear magnetic resonance through a completely different approach.[3] For these discoveries, Purcell and Bloch shared the nobel prize in physics in 1952. The practical use of NMR spectroscopy, however, didn’t gain momentum until the first commercially available spectrometer was introduced in 1958. Several groundbreaking discoveries were made in the field of NMR spectroscopy the following decades, for example the concept of two dimensional NMR spectroscopy[4] and the introduction of Fourier Transform NMR spectroscopy by Richard R. Ernst (nobel prize winner in chemistry in 1991). These developments together with increased availability of NMR spectrometers with superconducting magnets soon made NMR spectroscopy a routine method for analysis of chemical compounds. During the 1980’s, the area of Bio-NMR was developed by, amongst others, Kurt Wütrich, who was awarded the nobel prize in
Nuclear Magnetic resonance

chemistry in 2002. As the name implies, Bio-NMR includes the study of biological macromolecules such as proteins, DNA and RNA with NMR spectroscopy. Since then, alongside with the development of ever more powerful superconducting magnets, a rapid development and diversification of NMR spectroscopy have taken place. Today, NMR spectroscopy is an invaluable tool in both pharmaceutical industries and in academic research with a wide range of applications, such as chemical analysis of both liquid and solids, structure determinations of large biomolecules, screening for new drugs and metabonomics.

3.2. Basic NMR Theory

A nucleus can be viewed as a charged spherical object rotating about an axis, thus producing a small magnetic field. According to the rules of quantum mechanics, the spin angle momentum, \( \mathbf{P} \), of a nucleus is defined as

\[
P = \sqrt{I(I+1)}\hbar
\]  

(3.1)

Where \( I \) is the nuclear spin quantum number and \( \hbar \) is Planck's constant divided by \( 2\pi \). The value of \( I \) ranges from 0, 1/2, 1, 3/2, 2, ... up to 6 for different nuclei. Nuclei with \( I = 0 \) such as \(^{12}\text{C}\) and \(^{14}\text{N}\) are not detectable in NMR spectroscopy since they do not produce any magnetic field to interact with. The most commonly observed nuclei in NMR of organic molecules are \(^{1}\text{H}\), \(^{13}\text{C}\) and \(^{15}\text{N}\) which all have \( I = 1/2 \). In the following discussion, \( I = 1/2 \) is assumed.

\( \mathbf{P} \) is a vector and thus contains directional information. The orientation of this vector is also quantized. The spin angle momentum along an axis, for example z, is given by

\[
P_z = m\hbar
\]  

(3.2)

where \( m \) is the magnetic quantum number that can adopt 2I+1 values. For nuclei with \( I = 1/2 \), \( m \) has two values, \( \pm 1/2 \). All nuclei with \( I > 0 \) have a
magnetic dipole moment, $\mu$, associated with them. The magnetic dipole moment is directly proportional to $P$

$$\mu = \gamma P$$

(3.3)

where $\gamma$ is the gyromagnetic ratio of the observed nucleus. The gyromagnetic ratio is related to the internal structure of different nuclei but its origin is not fully understood and is beyond the scope of this text. The rather obscure quantum mechanical description of $P$ and $\mu$ can be rationalized by looking at a nucleus with $I = 1/2$ as a small bar magnet. When we expose such a nucleus to a static magnetic field with strength $B_0$, it will align itself relative to the magnetic field, just as a bar magnet or a compass needle. As the strength of the external magnetic field increases, so does the energy input needed to disrupt its alignment with the magnetic field. The same phenomenon is also observed when a stronger bar magnet is used, analogous to a nucleus with a larger gyromagnetic ratio. This is the basis for the fact that different nuclei have different resonance frequencies. As opposed to the analogy with a bar magnet, the energy difference between nuclei (often referred to as spins) aligned with and against the magnetic field, called the $\alpha$- and $\beta$-state, is very small, resulting in an almost equal population of spins in the $\alpha$- and $\beta$-state (Equation 3.4). This small population difference is the reason why NMR spectroscopy has a low sensitivity compared to other spectroscopic methods.

$$\frac{N_\beta}{N_\alpha} = e^{\Delta E/kT}$$

(3.4)

$$\Delta E = \gamma h B_0$$

(3.5)

$N_\beta$ and $N_\alpha$ are the populations of the $\alpha$ and $\beta$-state, respectively, and $k$ is the Boltzmann constant. This relationship is also illustrated in Figure 3.1a. The population difference, although small, produces a net magnetization, $M_z$, along the axis of $B_0$ (Figure 3.1b). The large gyromagnetic ratio of the proton and its natural abundance of almost 100% make it by far the most sensitive nucleus when studying organic
molecules with NMR spectroscopy. It is approximately ten times more sensitive than $^{13}$C and a hundred times more sensitive than $^{15}$N. Figure 1.1 also explains the quest for ever stronger magnetic fields since a larger ratio $N_\alpha/N_\beta$ results in a larger net magnetization and thereby a higher sensitivity of the experiment.

![Figure 3.1](image.png)

**Figure 3.1.** (a) The energy difference between the $\alpha$- and $\beta$-state is proportional to the external magnetic field $B_0$. (b) Alignment of the $\alpha$- and $\beta$-state of a nuclei with $m=1/2$ in a static magnetic field, $B_0$. The population difference between the $\alpha$- and $\beta$-state is greatly exaggerated to highlight the origin of the net magnetization, $M_z$. No net magnetization is observed in the x,y-plane since the individual magnetization vectors are evenly distributed in the x,y-plane.

As stated above, the individual nuclei in a sample will try to align themselves relative to an external magnetic field but due to their spin, they will start to rotate around the $B_0$-axis, much like a spinning top (Figure 3.1b). The precession rate of the nucleus is called its Larmor frequency and this angular velocity is given by

$$\omega_0 = \gamma B_0$$  \hspace{1cm} (3.6)

where $B_0$ is the strength of the applied magnetic field. In modern NMR instruments, the Larmor frequency for a proton ranges from 400-900 Mhz.
Nuclear Magnetic resonance

corresponding to magnetic field strengths of 9.4-21.1 T. Equation 3.6 suggests that all nuclei with the same gyromagnetic ratio would have exactly the same Larmor frequency. Fortunately, this is not the case since individual nuclei in a molecule always experience an effective magnetic field, $B_{\text{eff}}$, related to their chemical environment. The strength of $B_{\text{eff}}$ is given by

$$B_{\text{eff}} = (1 - \sigma)B_0$$

(3.7)

where $\sigma$ is the shielding constant. This constant is dependant on the electron density surrounding each proton and is thus the reason why different frequencies are observed for all protons with non-equivalent surroundings in a molecule. The local field that each nucleus experiences is also affected by the state of other magnetically active nuclei in close proximity in the covalent structure. This is called scalar- or J-coupling and results in splitting of the observed peak. J-coupling is usually observed only for spins separated by three or fewer bonds. As an example, the peak of an amide proton in an amino acid will appear as a doublet since its local field is influenced of whether the neighboring $\alpha$ proton exists in the $\alpha$ or the $\beta$-state (with the exception of Gly whose amide proton is coupled to two $\alpha$ protons). The frequency difference between the two peaks in the doublet is called the coupling constant. The most commonly observed J-couplings are between two protons two ($^2J$) or three ($^3J$) bonds apart in the structure. $^3J$-couplings are especially interesting in structure determinations of molecules because the coupling constant is dependent on the dihedral angle between the two protons, as discussed in chapter 4.2.

In order to create a signal, the bulk magnetization, $M_z$, has to be perturbed from its equilibrium. This can be achieved if the sample is subjected to a second magnetic field, $B_1$, applied perpendicular to $B_0$. The magnetization vector $M_z$ will then align itself along the new effective field $B_{\text{eff}}$ which is a sum of $B_0$ and $B_1$, thereby shifting part of the bulk magnetization from $M_z$ and creating transverse magnetization, $M_x$ and $M_y$. Rotation of the transverse magnetization vector will produce a detectable signal by inducing a current in the receiver coil (Figure 3.2). A static $B_1$-field would have to be of the same magnitude as $B_0$ which would be highly impractical. Instead, $B_1$ is applied as a rotating magnetic field with a
frequency close to the Larmor frequency of the nucleus studied. As discussed above, a molecule contain nuclei with many different frequencies, thus the sample must be irradiated with many different frequencies in order to create signals from all nuclei. Historically, this was accomplished by varying the frequency of the excitation radiation, similar to tuning in different radio stations with a radio transceiver, a method called continuous wave (CW) NMR spectroscopy. Modern NMR instruments, however, use the pulsed NMR technique, where the frequency range of the excitation needed to meet the resonance criteria of all nuclei is created by irradiation with a short, electromagnetic pulse with a frequency close to the Larmor frequency of the observed nucleus. As a consequence of Heisenberg’s uncertainty principle, a short pulse will contain a broad range of frequencies. If sufficiently short, such an electromagnetic pulse excites all nuclei within a few 100 kHz, more than enough to observe, for example, all protons with one pulse. The separation in Larmor frequencies of different nuclei, however, is large enough to prevent simultaneous excitation of several types of nuclei with one pulse. The oscillating signal observed in the receiver appears in the kHz range rather than the MHz range where the Larmor frequency is located. This can be rationalized by introducing the rotating frame concept where the rotating magnetic field $B_1$ is seen as a static field with the angular frequency $\omega_1$. The individual nuclei are then precessing relative to $\omega_1$ with an angular frequency, $\omega_{\text{eff}}$ of

$$\omega_{\text{eff}} = \sqrt{\omega_1^2 + (\omega_0 - \omega)^2}$$

(3.8)

where $\omega_0$ is absolute Larmor frequency according to equation 3.6 and $\omega$ is the effective Larmor frequency derived from equation 3.7. In other words, $\omega_{\text{eff}}$ is the difference between the frequency of the excitation pulse and the Larmor frequency of the individual spins.
In a simple one-pulse experiment, the signal, referred to as a free induction decay (FID), is a mixture of oscillating frequencies that contains the frequencies from all nuclei affected by the $B_1$-field. These signals, now in the time-domain (amplitudes as a function of time) are transferred to the frequency-domain by a Fourier transform resulting in an NMR spectrum. A consequence of equation 3.6 is that the same nucleus will have a different frequency depending on the strength on the magnetic field. This makes it difficult to compare spectra recorded on different spectrometers. Therefore, the chemical shifts are not given in Hz but in ppm (parts per million), a dimensionless unit determined by dividing the difference (in Hz) between a reference signal and the signal of interest with the Larmor frequency at the magnetic field strength used.

The relaxation of the signals is governed by two processes, $T_1$- and $T_2$-relaxation. $T_1$ or longitudinal relaxation describes the growth of the magnetization vector back to its original value along the $z$-axis. In other words, the rate by which the population difference between the $\alpha$- and $\beta$-states is restored. This process is for protons dominated by dipole-dipole relaxation[5] that gives rise to the NOE-effect, as described in chapter 3.3. The mechanism of $T_2$ or transverse relaxation involves loss of phase coherence in transverse magnetization vectors through mutual exchange of spin energies. The dominating cause for transverse relaxation
however, is often magnetic field inhomogeneities, resulting in slightly different precession frequencies in different parts of the sample.

3.3. Nuclear Overhauser Effect

Nuclei, or spins, in close vicinity to each other interact through the above mentioned dipole-dipole relaxation, resulting in population transfers between them. This can be explained by looking at the effect on one spin, called I, in a two-spin system when irradiating the other spin, S, selectively.

Figure 3.3. Energy level diagram of a two-spin system showing all transitions between the four energy levels.

A spin system containing two spin 1/2 nuclei close enough to be coupled via dipole-dipole coupling is comprised of four different energy levels corresponding to the $\alpha$- and $\beta$-states of the individual spins (Figure 3.3). The transitions that we can observe directly are the transitions between the $\alpha$- and $\beta$-state for the individual nuclei ($W_{1I}$ and $W_{1S}$). In presence of dipole-dipole coupling, two other interesting transitions, $W_0$ and $W_2$, appears. It can be realized that the intensity of the transitions $W_{1I}$ will be affected when altering the populations of spin S by selective irradiation of the $W_{1S}$ transition if magnetization is transferred through either $W_0$ or $W_2$. This is the basis for the nuclear Overhauser effect, NOE. The fractional change in signal intensity of spin I, $\eta_I$, when saturating spin S, i.e., equalizing the $\alpha$- and $\beta$-state populations of spin S, is defined as
\[ \eta_1 = \frac{(I - I_0)}{I_0} \quad (3.9) \]

I and \( I_0 \) are the signal intensities of I with and without saturation of spin S. In order for relaxation to occur through either \( W_0 \) or \( W_2 \), a local fluctuating field with the frequency corresponding to these energy differences must exist. A flux of such local fields appears continuously as a consequence of motions of the molecules that constitute the sample. The sign of the NOE depends on which relaxation pathway that dominates, \( W_0 \) or \( W_2 \). For small molecules with rapid tumbling rates, \( \tau_c \), the local magnetic fields induced by molecular motion will have high frequencies. Thus, \( W_2 \) will dominate, resulting in a negative NOE. Conversely, \( W_0 \) will dominate in large molecules with slow tumbling rates resulting in a positive NOE. In a transition region between positive and negative NOE, no or very small NOE’s are observed. This occurs for molecules with tumbling rates in the nanosecond time-scale, not unusual for peptides with a molecular weight of a few kDa. Acquiring the NOESY experiment on a spectrometer with a different magnetic field can solve this problem since the transition probabilities; or rather the proton frequency is dependent on \( B_0 \). Alternatively, rotating frame Overhauser effect spectroscopy, ROESY, can be used. In a ROESY experiment, a spin-lock is applied during the mixing time. The cross relaxation will thus take place in the presence of a weak RF-field instead of the strong static field, resulting in positive NOE for all molecules because the \( W_2 \)-transition will dominate regardless of tumbling rate. From the Solomon equation[6] that describes the relaxation of NMR signals, the cross-relaxation rate between spins I and S, \( \sigma_{IS} \), can be derived as

\[
\sigma_{IS} = \frac{\hbar^2 \gamma^2 r^{-6}}{10} \left( \frac{\mu_0}{4\pi} \right)^2 \left[ \frac{6}{1 + (\omega_1 + \omega_S)^2 \tau_c^{-2}} - \frac{1}{1 + (\omega_1 - \omega_S)^2 \tau_c^{-2}} \right] \quad (3.10)
\]

The take home message in this rather complex formula is that the cross-relaxation rate is proportional to the inverse sixth power of the distance between I and S, \( r \), making it possible to extract distance information from the measurement of cross peak volumes in a NOESY experiment.
4. Structure Calculations of Peptides

Knowledge of the three-dimensional structure of peptides, proteins and other biological macromolecules is essential in our understanding of the functions of biological systems at a molecular level. The sheer size of these molecules, however, makes such structure determinations a challenging task. NMR spectroscopy, conducted in solution, makes it possible not only to determine the three-dimensional structures of organic molecules but also to study dynamic processes such as enzymatic reactions and binding of ligands to their target molecules. Structure determination using NMR spectroscopy is an indirect method based mainly on a network of experimentally derived proton-proton distances from which a three-dimensional structure can be calculated. The algorithms used are targeted to minimize the violation of experimental restraints as well as keeping the covalent structure intact. The methodology of structure determinations of peptides using NMR spectroscopy was developed during the 1980’s.[7-9] Essentially the same methods are still used today, although developments in spectrometer hardware and experimental setups have enabled structure determinations of molecules and complexes exceeding 100 kDa.[10-12] An overview of the structure determination process is shown in Figure 4.1 and discussed further throughout this chapter.


**4.1. Assignment of Peptides**

A complete, or near complete assignment of the peptide is essential before any experimental restraints can be used in a structure determination. A well established strategy for the assignment of peptides have been used since the early 1980’s.[8] This strategy is based on a combination of two-dimensional experiments that utilizes both scalar couplings and NOE to give structural information about the peptide. Correlation Spectroscopy, COSY,[13] and Total Correlation Spectroscopy, TOCSY,[14] correlates protons in the peptide through their $^3$J-couplings. The difference between them is that in a COSY experiment, one cross peak is observed between two coupled protons and in a TOCSY experiment, cross peaks are observed between all protons in the same spin system (Figure 4.2). The TOCSY experiment is especially useful in identification of the amino acids in the peptide because no J-couplings spanning the amide bonds are present, making all amino acids isolated spin systems with specific cross peak patterns depending on the side chain structure. The NOESY experiment is an invaluable tool in sequential assignment of the peptide as well as in connecting separated spin systems.

---

**Figure 4.1.** Flowchart of the structure determination process using NMR spectroscopy. 

```
Sample
|↓|
NMR spectroscopy
|↓|
Resonance assignment
|↓|
Generation of restraints
|↓|
Structure calculation
|↓|
Analysis of structures
```
in side chains, such as Phe and Tyr. In a NOESY spectrum, cross peaks are observed between all protons in close proximity in space (up to approximately 5 Å). This makes it possible, in most cases (depending on the secondary structure of the peptide), to observe a cross peak between the amide proton of residue i+1 and the \( \alpha \) proton of residue i, thereby establishing the sequential order of residues i and i+1 (Figure 4.2).

**Figure 4.2.** Important cross peaks in an Ala-Ala dipeptide segment. Intra-residue connectivities observed in COSY and TOCSY spectra are marked with thin arrows and the sequential HN\(_{i+1}\)-H\( \alpha \_i \) connectivity observed in a NOESY experiment is marked with a broad arrow.

### 4.2. Distance and Angle Restraints

Distance restraints are by far the most important experimental restraints that can be derived from NMR spectroscopy. As mentioned in chapter 3.3, the distance between two protons can be calculated from cross peak volumes in a NOESY spectrum. In practice however, it is extremely difficult to derive an exact inter-proton distance, and in many cases one exact distance is not desired in structural determinations of peptides. The main reason for this is the inherent flexibility of short peptides. Rapid conversion between several conformers will inevitably give cross peak volumes and hence distance restraints biased towards the shortest distance present in the ensemble of conformers because of the \( r^6 \) dependence on the cross peak volume. Spin-diffusion is another issue that has to be taken into account. Spin-diffusion is the indirect cross relaxation between two protons via other protons in the surroundings (Figure 4.3). This can have a major impact on the cross peak volumes in a NOESY spectrum, especially for large molecules.
where $W_0$ is the dominating relaxation pathway. Cross peaks resulting from spin-diffusion can be detected by recording a series of NOESY spectra with different mixing times, $\tau_{\text{mix}}$. Cross peaks as a result of direct cross relaxation can be extrapolated to $\tau_{\text{mix}} = 0$ whereas cross peaks arising from spin-diffusion are delayed. The effect of spin-diffusion can be calculated by performing a relaxation matrix analysis in order to improve the accuracy of the distance restraints.[16, 17]

![Figure 4.3. Cross relaxations in a three-spin system, ABC.](image)

Fortunately, accurate distance restraints are not a prerequisite for a successful structure determination of a peptide since a large number of distance restraints are used. Instead, in most cases, the isolated spin pair approximation, ISPA, can be used. This method disregards any effects from spin-diffusion and is a scaling of all cross peak volumes using a volume from a spin pair with a known distance between them as a reference volume. By making the assumption that the rotational correlation time, $\tau_c$ is uniform throughout the molecule, the rather complex equation 3.10 can be simplified as

$$d_{ij} = d_{\text{ref}}(V_{\text{ref}}/V_{ij})^{1/6} \quad (4.1)$$

where $d_{ij}$ is the distance between protons $i$ and $j$ and $d_{\text{ref}}$ is the reference distance. $V_{ij}$ is the cross peak volume between protons $i$ and $j$ and $V_{\text{ref}}$ is the reference volume. Well-resolved methylene protons or H$_{\delta}$-H$_{\epsilon}$ cross peaks in the aromatic ring of tyrosine are common references in peptides. Because of the approximate value of $d_{ij}$ as well as a consequence of the inherent flexibility of peptides, $d_{ij}$ is not used as a fixed distance but rather as the basis for a distance interval that the protons are allowed to span without any energy penalties in the structure calculation (as
discussed in chapter 4.4). The size of the interval is a subjective choice that depends on the preference of the researcher and the quality of the spectra. Protons that cannot be treated individually, such as the equivalent protons in a methyl group or β-protons not assigned stereospecifically can be treated as pseudoatoms with radii encompassing the individual protons. When pseudoatoms are used, the upper limit of the distance restraint must be adjusted accordingly, resulting in loss of some precision in the structure calculation.[18] Therefore it is important to make as many stereospecific assignments as possible. This can be achieved by using local distance restraints together with scalar couplings[19] or by statistical analysis of structures resulting from preliminary structure calculations without any stereospecific information present.

The other major restraint useful in structure calculations of peptides that can be derived from NMR experiments is dihedral angles, most commonly backbone φ- and ψ-angles but also χ-angles in side chains. The relationship between dihedral angles and the $^3J$ coupling constant is described by the Karplus equation[20] that has the following general form:

$$^3J = A\cos^2\phi + B\cos\phi + C$$

(4.2)

where the coefficients A, B and C are empirically determined for different types of dihedral angles. The modified Karplus equation for the backbone φ-angle[21] and its solutions is shown in Figure 4.4. As seen in Figure 4.4, the Karplus equation has several solutions for some coupling constants, a fact that of course has implications when we want to use dihedral angle restraints in our structure calculations. A conservative use of only extreme $^3J$ coupling constants, leaving only one possible solution of equation 4.2 is one way to tackle this issue. Another way is to use dihedral angle restraints only as refining restraints after an initial structure calculation with only distance restraints. In the latter case, the solution of equation 4.2 that best fits to the initial structures is applied as a restraint. A more straightforward approach is the use of the coupling constants directly without the conversion to dihedral angles. This allow dihedral angles to assume values consistent with all solutions without energy penalty.[22] Coupling constants for the backbone φ-angle is relatively easy to measure from the splitting of the amide proton, either in a one-
dimensional experiment or in a double quantum filtered (DQF) COSY.[13] The coupling constant between $\alpha$– and $\beta$-protons, used for the determination of the $\chi_1$-angle is more difficult to determine because of the often complex cross peak patterns observed in a DQF-COSY. It is possible, however, to simulate the fine structure of such cross peaks using the programs SPHINX and LINSHA.[23, 24]

![Figure 4.4](image)

**Figure 4.4.** Solution to the Karplus equation parameterized for the $^3\!J$(HN-H$\alpha$) coupling.

### 4.3. Other Experimental Restraints

Hydrogen bonding restraints can be deduced from the temperature dependence of amide protons. The chemical shift of a hydrogen bonded amide proton is strongly influenced by the anisotropy contribution from the hydrogen bond acceptor. The bond lengthening of a hydrogen bond when the temperature is raised lessens this anisotropy contribution, leading to an upfield shift for the amide proton. The bond lengthening is more pronounced for hydrogen bonds involving the solvent than for intra-molecular hydrogen bonds. As a consequence, an amide proton in a protein or peptide with a small
A temperature coefficient (more positive than \(-4.5\) ppb/K) is thus considered to be involved in a hydrogen bond.[25] Care has to be taken in peptides, however, since small temperature coefficients can appear as a result of conformational exchange.[26] Predicted versus observed chemical shifts can also be used as restraints in structure calculations,[27] a method which is particularly useful in the determination of secondary structure elements.[28]

### 4.4. Structure Calculations

Distance geometry (DG) was the first method used to calculate structures of macromolecules with the aid of experimental restraints derived from NMR spectroscopy.[29, 30] Distance geometry is based on a distance matrix that contains all known distances within the molecule, covalent information as well as experimental restraints. An embedding algorithm can then convert the distance matrix to cartesian space. This is a simple and elegant method for the generation of structures but is however, associated with some drawbacks. One problem with distance geometry is the fact that some distances are missing or approximate. These values are subsequently chosen randomly for each calculated structure as long as no other distances are violated, resulting in structures that in most cases have the correct fold but can be severely distorted and need to be subjected to extensive energy minimization. Moreover, structures calculated with distance geometry can yield similar structures trapped in local energy minima even after extensive sampling of the given distances.[31, 32] Distance geometry can be useful, for example, as a tool to generate structures that can be used as starting points in molecular dynamics structure calculations.

The dominating method to calculate NMR structures of macromolecules today is molecular dynamics, MD, in cartesian space. MD calculations are based on integration of Newton’s equation of motion in order to get a trajectory of the molecular motion. The motions of a large molecule is very complex and the algorithms (force fields) used for NMR-based structure determination do not strive to accurately simulate all forces involved. Instead, they are designed to generate structures...
consistent with the experimental restraints as well as the covalent structure. Newton’s equation of motion (Equation 4.3) relates the positions of the individual atoms at a given time to the forces that act upon them.

\[ m_i \frac{d^2 r_i}{dt^2} = F_i \quad (i = 1, \ldots, n) \]  

(4.3)
m_i is the mass of atom i, r_i is the position of atom i at time t, F_i is the force on atom i at time t and n is the number of atoms. F_i is given by the negative gradient of the potential energy function, the force field. One of the most commonly used force fields is X-PLOR[33] that contains the following energy terms:

\[
E = \sum_{bonds} k_b (b - b_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} k_\phi (1 + \cos(n\phi + \delta)) \\
+ \sum_{impropers} k_\phi (\phi - \delta)^2 + \sum_{VDW} k_{VDW} (\max(0, sR_{min})^2 - R^2)^2 \\
+ E_{NOE} + E_{dih}
\]

(4.4)

Where b and b_0 are the actual and ideal bond lengths, \(\theta\) and \(\theta_0\) the actual and ideal angles, \(\phi\) the actual dihedral or improper angle, \(\delta\) an offset value of the dihedral or improper angle, s a scaling factor, \(R_{min}\) is the value where the van der Waals potential has its minimum and \(k_b\), \(k_\theta\), \(k_\phi\), and \(k_{VDW}\) are the force constants for each energy term. The potential energies for the experimental distance and angle restraints, \(E_{NOE}\) and \(E_{dih}\), are defined as square well potentials, allowing the distances and dihedral angles to vary within a specified interval without energy penalty (Equations 4.5 and 4.6).

\[
E_{NOE} = \sum \begin{cases} 
k_{NOE} (d_{ij} - u_{ij})^2 & \text{if } d_{ij} > u_{ij} \\
0 & \text{if } l_{ij} \leq d_{ij} \leq u_{ij} \\
k_{NOE} (l_{ij} - d_{ij})^2 & \text{if } d_{ij} > u_{ij} \end{cases}
\]

(4.5)

\[
E_{dih} = \sum \begin{cases} 
k_{dih} (\phi_{ij} - \phi^{u}_{ij})^2 & \text{if } \phi_{ij} > \phi^{u}_{ij} \\
0 & \text{if } \phi_{ij} \leq \phi_{ij} \leq \phi^{u}_{ij} \\
k_{dih} (\phi^{l}_{ij} - \phi_{ij})^2 & \text{if } \phi_{ij} > \phi^{u}_{ij} \end{cases}
\]

(4.6)
k\text{NOE} and k\text{dih} are force constants for the distance and angle restraints, d_{ij} and \phi_{ij} are the distance and dihedral angle between atoms i and j at a given time, u_{ij} and \phi^{u}_{ij} are the upper limits for the restraints and l_{ij} and \phi^{l}_{ij} are the lower distance and angle limits.

In order to get a better sampling of the conformational space and to avoid local energy minima the temperature of the system can be raised. This enables transitions across high-energy barriers in the molecule. The temperature can be controlled by scaling the velocities for each atom during the MD simulation or by coupling to a heat bath.[34] The temperature is lowered throughout the simulation in order to freeze out low-energy conformations, a procedure called simulated annealing, SA. A typical simulated annealing protocol, as used in papers I and III consist of four parts: (i) initial conjugate gradient energy minimization; (ii) MD simulation at a high temperature (1000 K) for several picoseconds; (iii) cooling the system from 1000K to 100 K; (iv) a final energy conjugate gradient minimization after which the atomic coordinates are saved. Although simulated annealing is a good method to generate structures located in the global energy minimum, it cannot be ruled out that the resulting structures are influenced on the initial conformation of the molecule. Therefore, the MD simulation is often repeated with different starting structures. Structure calculations are often iterative processes (as indicated in Figure 4.1) where results from initial MD simulations are used to evaluate the experimental restraints in order to detect inconsistencies such as wrongly assigned resonances.

Analysis of the resulting structure ensemble is an important part of any structure calculation since it gives a measure of the validity and precision of the structures. In most cases, violation of experimental restraints by a certain amount or a total energy limit are used as rejection criteria on individual structures in the ensemble of calculated structures. Root mean square deviation, rmsd, of the atoms is used as a measure of the structural variation within the ensemble of conformers. In cases where violations of distance restraints occurs as a consequence of conformational exchange, time-averaged distance restraints can be used.[15] As the name implies, time-averaged restraints do not have to be fulfilled at any given time during a MD simulation but over a defined time of the molecular trajectory.
5. Analogues of Biologically Active Peptides

5.1. Turn Mimetics and Amide bond Isosteres

Peptide hormones regulate a large variety of functions in the human body and the understanding of how they interact with their respective receptors is of great interest in the development of new drugs. Therefore, efforts to determine the relationship between conformation and activity in biologically active peptides have been a challenging task for decades. The use of the peptides themselves as drugs is limited by their poor pharmacokinetic properties such as low uptake, rapid degradation and secretion.[35] In order to overcome these issues, there is an ongoing search for peptidomimetics, modified peptides or non-peptide molecules that can act as agonists or antagonists at the same receptor as the endogenous peptides. The fact that most hormone receptors are membrane-bound proteins has hampered the elucidation of the bioactive conformations of peptide hormones. Such receptors are usually difficult to obtain in large amounts and structure determination through crystallization or by using NMR spectroscopy is difficult, or impossible. Thus only a few structures of membrane-bound proteins are available.

Consequently, the focus has been directed at modification of the endogenous peptide hormones in order to get information about the bioactive conformation and to pinpoint structural features of importance for the biological activity. A multitude of different approaches have been investigated when it comes to designing peptidomimetics, including backbone modifications and mimetics of secondary structures such as turns, helices and β-sheets.[36-38] (selected reviews) Besides maximizing the potency of the substance, the aim of a peptidomimetic could be to
increase the resistance towards enzymatic degradation by, for example, modifying the backbone, or to reduce the flexibility of the peptide, thereby decreasing the entropy-cost upon receptor binding.

Turns, defined as a region where the backbone changes its overall direction, constitute an important class of secondary structure in both peptides and proteins. The most common types of turns are γ- [39] and β-turns,[40] figure 5.1.

Figure 5.1. Schematic representation of γ- and β-turns. Rotations around the marked bonds give rise to the different sub-types of turns summarized in table 5.1.

In a γ-turn, the backbone changes direction over three residues and a hydrogen-bond is formed between the carbonyl oxygen of residue i and the amide hydrogen of residue i+2, thus forming a pseudo 7-membered ring. γ-Turns are divided into two groups, classical and inverse, depending on the backbone angles of residue i+1, Table 5.1. In a β-turn, the backbone reverses its direction over four residues and a hydrogen bond between the carbonyl oxygen of residue i and the amide proton of residue i+3 is often formed, resulting in a pseudo 10-membered ring. A variety of different classes and subclasses of β-turns exist, depending on the backbone angles of residues i+1 and i+2, Table 5.1. Several peptide hormones have been found to contain γ- or β-turns, including oxytocin[41] which induces milk production and labour in mammals, vasopressin[42] which controls the reabsorption of water in the kidneys, Leu-enkephalin[43] which has an analgesic effect and angiotensin which is involved in regulation of blood-pressure.[44]
### Table 5.1. Backbone dihedral angles defining the different types of γ- and β-turns.

<table>
<thead>
<tr>
<th>Turn</th>
<th>Backbone angle (deg)</th>
<th>i+1</th>
<th>i+2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>φ</td>
<td>ψ</td>
<td>φ</td>
</tr>
<tr>
<td>Type  Iβ</td>
<td>-60</td>
<td>-30</td>
<td>-90</td>
</tr>
<tr>
<td>Type  I’β</td>
<td>60</td>
<td>30</td>
<td>90</td>
</tr>
<tr>
<td>Type  IIβ</td>
<td>-60</td>
<td>120</td>
<td>80</td>
</tr>
<tr>
<td>Type  II’β</td>
<td>60</td>
<td>-120</td>
<td>-80</td>
</tr>
<tr>
<td>Type  IIIβ</td>
<td>-60</td>
<td>-30</td>
<td>-60</td>
</tr>
<tr>
<td>Type  III’β</td>
<td>60</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>Type  VIaβ</td>
<td>-60</td>
<td>120</td>
<td>-90</td>
</tr>
<tr>
<td>Type  VIbβ</td>
<td>-120</td>
<td>120</td>
<td>-60</td>
</tr>
<tr>
<td>γ-turn</td>
<td>70-85</td>
<td>-60-(-70)</td>
<td></td>
</tr>
<tr>
<td>Inverse γ-turn</td>
<td>-70-(-85)</td>
<td>60-70</td>
<td></td>
</tr>
</tbody>
</table>

In order to investigate if the turns found in solution reflect the bioactive conformation when the hormones are bound to their respective receptors we have investigated the structure and activity of desmopressin containing a γ-turn mimetic as well as a methylene ether isostere (paper I), and also Leu-enkephalin with a β-turn incorporated (paper III). A new synthetic route towards the methylene ether isosteres \( \text{Xxx}Ψ[\text{CH}_2\text{O}]\text{Ala/Gly} \) was also developed (paper II).

### 5.2. Studies of Desmopressin

#### 5.2.1. Biological Background

Desmopressin is a successful example of how modifications of an endogenous peptide hormone can result in a compound suitable for pharmaceutical use. This peptide originates from the neurohypophyseal peptide hormone vasopressin, a cyclic nonamer, Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly (cf. 5.1 in Figure 5.2), with a disulfide bridge linking Cys1 and Cys6 and a three-residue acyclic tail. Vasopressin exerts a number of physiological activities in our bodies,
Analogues of Biologically Active Peptides

most notably blood pressure regulation and antidiuretic activity but it also influences behavioural processes such as memory, anxiety and depression. These actions are triggered by the interaction of three different G-protein coupled receptors, GPCRs, the V₁a, V₁b and V₂ receptors.

Desmopressin is the result of two modifications of vasopressin, deamination of Cys₁, to give mercaptopropionic acid (Mpa), and replacement of Arg₈ with D-Arg, Figure 5.2.

\[ \text{Cys - Tyr - Phe - Gln - Asn - Cys - Pro - Arg - Gly - NH₂} \]

Vasopressin (5.1)

\[ \text{Mpa - Tyr - Phe - Gln - Asn - Cys - Pro - D-Arg - Gly - NH₂} \]

Desmopressin (5.2)

Figure 5.2. The structures of vasopressin, 5.1, and desmopressin, 5.2. Residues in bold highlights the structural differences between the two peptides.

Desmopressin is more selective towards the V₂ receptor, responsible for the reabsorption of water in the kidneys, than vasopressin and is therefore used in treatment of, for example, diabetes insipidus[45] and nocturia[46].

NMR has been used in several conformational studies of desmopressin both in aqueous solution[47] and in aqueous solution containing 20% and 70% trifluoroethanol (TFE).[48, 49] In the presence of TFE, a type II β-turn was found with residues Pro₇ and D-Arg₈ in the i+1 and i+2 positions. A type I β-turn with residues Gln₄ and Asn₅ in the i+1 and i+2 positions was also found in 2 out of 5 calculated structure families.[48] Both turns were found to be open, that is, no stabilizing hydrogen bonds were observed. A similar conformation was found when desmopressin bound to the carrier protein neurophysin-II was investigated with a transfer NOE experiment.[50] In aqueous solution, on the other hand, an inverse γ-turn was found, centred on Gln₄. The same type II β-turn in the acyclic tail was observed as found in the presence of TFE. The
conformational differences between these studies can be rationalized as an effect of TFE which is known to destabilize solvent-exposed amide functions.[51] All in all, it can be concluded, not surprisingly, that desmopressin is a flexible peptide, and that its conformation is dependent on the surrounding environment. Therefore, no definite conclusions about the bioactive conformation when bound to the V2 receptor can be drawn from the studies mentioned above. In order to investigate if the inverse \( \gamma \)-turn found in desmopressin in aqueous solution is present in its bioactive conformation, an inverse \( \gamma \)-turn mimetic was designed and incorporated into desmopressin (cf. 5.3, Figure 5.3). This mimetic consists of a methyl substituted morpholine-3-one ring. The hydrogen bond between the carbonyl oxygen of Phe3 and the amide proton of Asn5 that was observed in aqueous solution is replaced by a methylene bridge and the amide bond between Phe3 and Gln4 is replaced with a \( \Psi[CH_2O] \) isostere.[52] The methyl group in the C-2 position of the morpholine-3-one ring mimics an alanine side chain, replacing Gln4; a substitution that only has a minor effect on the biological activity.[53] This mimetic serves the dual purpose of locking the \( \gamma \)-turn conformation as well as modifying the amide bond following Phe3; an amide bond otherwise susceptible to enzymatic cleavage by chymotrypsin. The stereochemistry at the C-6 position determines whether the mimetic resembles an inverse- or a classical \( \gamma \)-turn (cf. 5.3 and 5.4, respectively, Figure 5.3). To be able to evaluate the effect of the methylene ether isostere replacement separately, a dipeptide mimetic, Phe\( \Psi[CH_2O] \)Ala was synthesized and incorporated into desmopressin. This mimetic is interesting also since it alters the hydrogen bonding possibilities of the peptide without a significant change in the Phe3 C\( \alpha \) – Ala4 C\( \alpha \) distance. The flexibility is also increased since the rotational barrier in a methylene ether isostere is 2.7 kcal/mol compared to 21 kcal/mol in an amide bond.[54] A total of three desmopressin analogues have thus been synthesized, one containing an inverse \( \gamma \)-turn mimetic, 5.3, a second containing a classical \( \gamma \)-turn mimetic, 5.4, and a third containing the Phe\( \Psi[CH_2O] \)Ala dipeptide mimetic, 5.5, figure 5.3.
Analogues of Biologically Active Peptides

Figure 5.3. Three desmopressin analogues containing an inverse γ-turn mimetic, 5.3, a classical γ-turn mimetic, 5.4, and a methylene ether isostere, 5.5.

The synthesis of desmopressin analogues 5.3 and 5.4 was developed by Dr Kay Brickmann and Dr. ZhongQing Yuan. A retrosynthetic analysis showed that the morpholine-3-one scaffold in 5.6 and 5.7 could be synthesized from azido-epoxides 5.11 or 5.12 and the protected amino alcohol 5.13 together with the bromo acid 5.10 (Scheme 5.1). The stereochemistry of the azido-epoxides can be varied depending on the wanted stereochemistry at the C-6 position in the morpholine-3-one ring. The inverse and classical γ-turn mimetics, 5.6 and 5.7, containing Phe, Ala and Asn in positions i, i+1 and i+2 were prepared in 7-step syntheses with overall yields of 27% and 32%, respectively, from epoxides 5.11 and 5.12.[52] These γ-turn mimetics, with side chains corresponding to that of desmopressin were used in solid-phase peptide synthesis of peptides 5.3 and 5.4, respectively.
Analogues of Biologically Active Peptides

Scheme 5.1. Retrosynthetic analysis of γ-turn mimetics 5.6 and 5.7.

5.2.2. Stereoselective Synthesis of Ψ[CH₂O] Pseudodipeptides

In order to synthesize 5.5, we needed a stereoselective route towards the PheΨ[CH₂O]Ala pseudodipeptide. Several routes towards these types of methylene ether isosteres have been reported previously. These include Williamson’s ether synthesis between amino alcohols and α-bromo substituted Gly or Ala[54, 55], rearrangement via a morpholine-3-one ring of a dipeptide with the C-terminal carboxylic acid reduced to an alcohol[55, 56] and aziridine ring opening with an alcohol.[57] Both the Williamson’s ether synthesis and the intramolecular rearrangement mentioned above were investigated in our attempts to synthesize a PheΨ[CH₂O]Ala pseudodipeptide. Unfortunately, the formation of diastereomers which were difficult to separate occurred in both cases.[58] Introduction of two modifications in the Williamson’s ether synthesis route proved to be successful, however. It was reasoned that the Fmoc-protecting group used in the first attempts could be the reason why the products were difficult to purify and handle. Therefore, the amine of the first residue in the pseudodipeptide was masked as an azide (Scheme 5.2).
The azide serves as a sterically unhindered protecting group of the amine functionality and can easily be reduced after incorporation of this pseudodipeptide building block into the peptide.[59] The N≡N stretch of the azide is easily distinguished using IR spectroscopy and the reduction can subsequently be monitored with IR spectroscopy on a few beads of the resin used in the solid phase peptide synthesis. The other modification was to use chlorine instead of bromine as the leaving group in the substitution, (R)-chloropropionic acid was thus used instead of (R)-bromopropionic acid. In order to test the robustness of this synthesis, we also prepared an IleΨ[CH₂O]Ala pseudodipeptide that corresponds to the first two residues of the CII260-267 glycopeptide. This is a part of type II collagen which plays a central role in eliciting disease in mouse models for rheumatoid arthritis[60] and possibly also in humans.[61]

The azide functionality in 5.16 and 5.17 was the result of a copper-catalyzed diazo transfer reaction[62] performed on the amino...
alcohol derivatives of Phe (5.14) and Ile (5.15), a reaction that worked very well for the preparation of 5.17 (90%) and moderately well for 5.16 (68%). Alkylation of 5.16 and 5.17 with (R)-2-chloropropionic acid under basic conditions furnished the methylene ether isosteres 5.18 and 5.19 in 48% and 63% yields, respectively. The exchange of (R)-2-bromopropionic acid for (R)-2-chloropropionic acid dramatically improved the stereoselectivity of this reaction. Only trace amounts, 1-3% according to $^1$H and $^{13}$C NMR spectroscopy, was found of the unwanted diastereomers PheΨ[CH$_2$O]D-Ala and IleΨ[CH$_2$O]D-Ala compared to approximately 25% when using bromine as leaving group. Desmopressin analogue 5.5 and the type II collagen glycopeptide CII260-267 (5.20), were prepared by solid-phase peptide synthesis in 23% and 28% yields, respectively. The azide functionalities of 5.18 and 5.19 were reduced, in quantative yield according to IR, during the peptide synthesis, with tin(II) chloride in the presence of thiophenol and triethylamine.[59] The yield in the peptide coupling following the reduction was impaired unless special care was taken to wash the resin with 20% piperidine in DMF after the reduction, probably as a result of complex formation between the liberated amine functionality and tin-salts.

5.2.3. Structure Determination and Biological Evaluation of Desmopressin Analogues

The conformations of desmopressin analogues 5.3 and 5.5 were studied with NMR spectroscopy and their biological activity was evaluated in order to gain insight towards the bioactive conformation of desmopressin. The assignment of the peptides were done in aqueous solution at 5 °C and pH 6.7 using conventional peptide assignment strategy discussed in chapter 4.1[8] based on a combination of DQF-COSY[13], TOCSY[14], NOESY[63] and gradient-enhanced $^{13}$C-HSQC experiments.[64] In addition, a ROESY[65] experiment was acquired for peptide 5.5. All spectra were recorded on a Bruker DRX 600MHz spectrometer equipped with a triple-resonance probe with x, y and z-gradient coils.

In addition to NOE cross-peak intensities, amide proton temperature coefficients, and also $^3$J$_{NH}$ and $^3$J$_{HetH}$ coupling constants...
Analogues of Biologically Active Peptides

were measured in order to determine hydrogen bond- or dihedral angle restraints. The only amide proton temperature coefficient with a value more positive than -3 ppb/K was that of Phe3 in peptide 5.5 (-1.5 ppb/K), a value close to that found in desmopressin (-2.4 ppb/K).[47] This similarity is reflected by a structural similarity as revealed in the structure calculation. In desmopressin, the temperature coefficient of Phe3 was argued to be a result of shielding of the amide proton from the solvent rather than the involvement in hydrogen bonding. In general, care has to be taken in the use of hydrogen bonds as restraints since small temperature coefficients also can reflect the loss of secondary structure.[26] Therefore, no hydrogen bond restraints were used in the structure calculations for 5.5. The $^3\text{J}_{\text{NHe}}$ of Phe3 in 5.3, 9.6 Hz was converted to a $\phi$ dihedral angle restraint of -109±30°. All other $^3\text{J}_{\text{NHe}}$ for both 5.3 and 5.5 were in the region of conformational averaging[66] and thus not used as restraints. $\chi^1$ torsion angle restraints were estimated through simulation of well-resolved $H_{\alpha}$-$H_{\beta}$ cross peaks with the programs SPHINX and LINSHA.[24] In this way, $\chi^1$ torsion angle restraints for Phe3 and Cys6 in peptide 5.3 were calculated to $-50\pm30$° and $-70\pm30$°, respectively. Due to overlap or degeneracy of $\beta$-protons, no torsion angle restraints were determined for peptide 5.5. Since the Karplus relation gives multiple solutions in the conversion between coupling constants and dihedral angle, the restraints mentioned above are from the solution compatible with initial structure calculations without any dihedral angle restraints.

In total 74 distance restraints were collected from the NOESY spectrum of peptide 5.3 while 55 distance restraints were derived from the ROESY spectrum of peptide 5.5. The conversion between NOESY cross peak intensities and distance restraints was made using the isolated spin-pair approximation with the mean value of the cross peak volumes between the geminal $\beta$-methylene of Phe3 and Cys6 used as reference volume for peptide 5.3. Because of overlap, only the cross peak volume between the $\beta$-protons of Phe3 could be used as reference for peptide 5.5. The reference distance was set to 1.77 Å. The restraints were used in ab initio simulated annealing calculations with the X-PLOR force field.[33] The presence of conformational exchange became apparent in initial structure calculations of peptide 5.3 since contradicting distance restraints resulted in high-energy structures with numerous distance restraint
violations. A closer inspection of the resulting structures revealed the presence of two different structure families, named A and B, with complementary NOE violations. Thus, it was concluded that the high-energy structures was the result of conformational exchange between the two structure families. As a consequence, the list of distance restraints was split in two subsets, each containing the distance restraints found to be valid for structure families A and B, respectively. This resulted in well-defined low-energy conformations without any distance restraint violations larger than 0.3 Å with heavy atom rmsd values for residues 1-6 of 0.71 and 0.63 Å for families A and B (Figure 5.4a and b).

**Figure 5.4.** Superimposition of 10 random structures from the ensembles of accepted structures of a) structure family A of peptide 5.3, b) structure family B of peptide 5.3 and c) peptide 5.5.
The conformational exchange between families A and B involves a flip of the morpholinone-ring relative to the plane of the macrocyclic ring coupled to a relocation of the Asn5 side chain. The broad resonance of Asn5 H$_{\alpha}$ that puzzled us initially fits well with this result. Chemical shift calculations with Gaussian98 reveals that it experiences a large difference in chemical shift in structural families A and B, respectively, mainly due to the difference in distance to the carbonyl oxygen in the morpholinone-ring. This results in a situation where Asn5 H$_{\alpha}$ is near coalescence while the other protons are in the fast exchange regime. This also implies that the exchange rate is in the millisecond time-scale which is too slow to simulate with unrestrained molecular dynamic simulations that otherwise could have been used to confirm the existence of this exchange process. The morpholin-3-one ring keeps its intended conformation with the Ala4 side chain in a pseudo-equatorial position in both structure families.

Peptide 5.5 showed no signs of conformational exchange of the same magnitude as peptide 5.3. Instead, the backbone adopted one well-defined conformation with a heavy atom rmsd of 0.56 Å for residues 1-6 (Figure 5.4c). The fold of the macrocyclic ring of peptide 5.5 resembles that of structure family A of peptide 5.3 with the side chain of Asn5 pointing towards the middle of the ring. The three-residue tail shows considerable flexibility in both peptides although observed hydrogen bonds imply that a turn is preferred. Hydrogen bonds between Cys6 O and Gly9 HN and Cys6 O and D-Arg8 HN were observed in the ensemble of accepted structures of both peptides. The former was more frequent in peptide 5.5 than in peptide 5.3 (50% occurrence in 5.5, 13% in 5.3 A and 29% in 5.3 B) while the latter was found in all structures of peptide 5.5 and in 36% of the ensemble of peptide 5.3. Consequently, the tail seems to have a preference for either a $\gamma$- or a $\beta$-turn.

Biological testing of the antidiuretic effect, that is, the agonist activity at the V$_2$ receptor, of peptides 5.3, 5.4 and 5.5 was performed in a rat-model, table 5.2. As can be seen, the two desmopressin analogues with the incorporated $\gamma$-turn mimetics, 5.3 and 5.4, as well as the methylene ether isostere containing analogue, 5.5, have agonist activities well below that of desmopressin and [Ala$^4$]desmopressin, which is the most relevant reference substance. Virtually no agonist activity at all could be detected.
Analogues of Biologically Active Peptides

for peptide 5.3 and 5.4 while the more flexible analogue 5.5 shows roughly a 1000-fold decrease in activity compared to [Ala4]desmopressin.

Table 5.2. \(^{a}\)Number of experiments. \(^{b}\)The antidiuretic activity was calculated using an international vasopressin standard (561 IU/mg). Values are mean (±SEM).

<table>
<thead>
<tr>
<th>peptide</th>
<th>n (^{a})</th>
<th>agonist activity(^{b}) (IU/mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>vasopressin 5.1</td>
<td>-</td>
<td>561</td>
</tr>
<tr>
<td>desmopressin 5.2</td>
<td>-</td>
<td>1000</td>
</tr>
<tr>
<td>[Ala4]desmopressin</td>
<td></td>
<td>645</td>
</tr>
<tr>
<td>5.3</td>
<td>3</td>
<td>&lt;2⋅10(^{-3})</td>
</tr>
<tr>
<td>5.4</td>
<td>6</td>
<td>2.6⋅10(^{-3}) (±3.7⋅10(^{-4}))</td>
</tr>
<tr>
<td>5.5</td>
<td>7</td>
<td>0.56 (±7.2⋅10(^{-2}))</td>
</tr>
</tbody>
</table>

Comparisons of the calculated structures of 5.3 and 5.5 to that of desmopressin in aqueous solution give some hints as to why such a dramatic effect in the activity is observed. As can be seen in figure 5.5, both structure families A and B of peptide 5.3 differ in their backbone conformations compared to desmopressin (backbone rmsd-values of residues 1-6 of 1.31 and 1.80 Å, respectively). The side chains of Tyr2 and Phe3 are positioned roughly at the same position as in desmopressin in both families while the positions of the side chains of Ala4 and Asn5 have similar positions as desmopressin in family A but differ quite dramatically from desmopressin in family B. Because Asn5 is critical to the antidiuretic effect of desmopressin\(^{[67]}\), this conformational exchange could to some extent be a cause of the reduced activity for peptide 5.3. Peptide 5.5 on the other hand has a very similar backbone fold as desmopressin with a backbone rmsd for residues 1-6 of only 0.66 Å. The slightly different position on the side chain of Asn5 is the only noticeable difference, a difference probably caused by the fact that the hydrogen bond between Phe3 CO and Asn5 HN cannot be formed in peptide 5.5. The dramatic drop in activity for peptide 5.5 was rather surprising considering this relatively small change in conformation. Thus, it can be concluded that the Phe3-Ala4 amide bond is an important structural feature in desmopressin, most likely as a hydrogen bond donor to the receptor or in intramolecularly when bound to the receptor.
Figure 5.5. Comparison of the lowest energy structure found in aqueous solution for desmopressin with the lowest energy structures from a) family A of peptide 5.3, b) family B of peptide 5.3 and c) peptide 5.5. Desmopressin is represented in grey. For clarity, only residues 1-6 are shown. The backbone atoms (N, Ca, C) of residues 1-6 were used for the superimposition and the rmsd calculations.
5.3. Studies of Leu-enkephalin

5.3.1. Biological Background

Development of analgesics is a pharmaceutical area of great interest, both historically and presently. The most well-known analgesic is probably the alkaloid morphine that has been used in its raw form, opium, as early as 3400 BC in Mesopotamia, and later, from the 19th century and onward, as an isolated compound. A severe drawback of the use of morphine and related compounds such as heroin is the development of tolerance and addiction. In the American civil war alone, some 400,000 people developed morphine addiction. Therefore, the development of new, more potent, analgesics with fewer side effects has been an ever-ongoing process throughout the 20th century.

Leu-enkephalin is an endogenous pentapeptide, Tyr-Gly-Gly-Phe-Leu (c.f. 5.21 in Figure 5.6), with affinity for the opioid receptors, the same receptors that morphine acts on. The opioid receptors are divided in three sub-classes, δ, μ and κ, all of which are transmembrane G-protein coupled receptors. The similar effects of morphine and Leu-enkephalin has resulted in large efforts to determine the bioactive conformation of Leu-enkephalin in order to produce more potent analgesics.[68] (selected review) This has proven to be quite complex since the peptide exhibits different binding modes for the different receptor subtypes,[69] a fact that also can be advantageous when a specific receptor is targeted. Like all short peptides, Leu-enkephalin is flexible and can populate a multitude of conformations, amongst which a β-turn with the two glycine residues in position i+1 and i+2 has been found both in solution[70] and in the crystal structure.[71] Based on these results, a β-turn mimetic mimicking the observed 1→4 β-turn has been synthesized an incorporated into Leu-enkephalin. This analogue has been subjected to structure determination and biological evaluation.
5.3.2. Design and synthesis of a 10-membered ring β-turn mimetic

As discussed in chapter 5.1, a β-turn is often stabilized by a hydrogen bond between the carbonyl oxygen of residue i and the amide proton of residue i+3. The distance between the carbonyl carbon in residue i and the amide nitrogen in residue i+3 varies between 3.5-3.8 Å for various types of β-turns, a distance that corresponds well to that of an ethylene group. Replacement of the hydrogen bond with an ethylene group results in a 10-membered ring that besides the obvious property of stabilizing the β-turn also features modified amide bonds between i and i+1 as well as i+3 and i+4, which could increase the stability of the molecule in vivo. Such a β-turn mimetic was synthesized as a mimetic of Leu-enkephalin (Figure 5.6).

The synthesis of the β-turn mimetic was developed by David Blomberg. A retrosynthetic analysis of compound 5.22 and its protected precursor 5.23 revealed that it could be prepared from β-azido alcohol 5.24, α-bromo acetic acid 5.25, Fmoc-protected glycine (5.26), and the Phe-Leu dipeptide 5.27, (Scheme 5.3). The double bond in the β-azido alcohol 5.24 serves as an aldehyde precursor used in the formation of the ethylene
bridge through a reductive amination with dipeptide 5.27. The 10-membered ring in 5.23 was formed by intramolecular ring-closure between the acetic acid and glycine moieties, after attachment of these building blocks to the product obtained from condensation of 5.24 and 5.27. In total, 5.22 was prepared in 3.2% over 15 steps.

Scheme 5.3. Retrosynthetic analysis of a β-turn mimetic containing Leu-enkephalin, 5.22.
5.3.3. Structure Determination and Biological Evaluation of a $\beta$-turn Containing Leu-enkephalin Analogue

The relevance of a $1\rightarrow 4$ $\beta$-turn in the bioactive conformation of Leu-enkephalin was investigated through conformational analysis on the protected Leu-enkephalin mimetic 5.23 and biological evaluation of the unprotected Leu-enkephalin mimetic 5.22. It would have been preferable to determine the structure of 5.22 as well but unfortunately, it turned out that conformational exchange between several conformers in the slow or intermediate time scale made such a conformational analysis based on NMR spectroscopy difficult. This conformational exchange revealed itself as broad or multiple resonances for a large portion of the protons in the compound as well as the occurrence of cross peaks of both negative and positive signs in the ROESY spectrum. This problem could not be circumvented although several solvents (DMSO:H$_2$O, TFE:H$_2$O and MeOH) and temperatures were tested. Finally, we found that the protected mimetic 5.23, scheme 5.3, gave spectra of sufficient quality for conformational analysis in MeOH at $-70$ °C (cooled with liquid nitrogen). The azide-group that masks the amide of Tyr1 and the methyl ester protected C-terminal should not alter the conformation of the $\beta$-turn mimetic and 5.23 could thus serve as a template for the conformation of 5.22. Because of the rather extreme temperature, it was not surprising that several of the NOESY cross peaks observed in the original NOESY spectrum with a mixing time of 300 ms were the result of spin-diffusion instead of direct dipole-dipole relaxation. This became evident through the acquisition of six NOESY spectra with the mixing time varied from 30 to 300 ms. In total 29 distance restraints were derived from the NOESY spectrum with a mixing time of 30 ms in the same way as described in section 5.2.3 with the cross peak volume of Gly$_{2a1}$-Gly$_{2a2}$ used as the reference volume. These restraints were divided into short and long distances with the upper limits set to 4 and 5 Å, respectively. Structure calculation with the X-PLOR force field resulted in an ensemble of structures with a very well defined backbone conformation in the turn mimetic part. Since the stereochemical assignment of the methylene protons in the bridge connecting Tyr1 with Phe4 was based solely on their respective NOEs, a second structure calculation where the two methylene groups were treated as pseudoatoms.
was performed. The resulting structures did not differ significantly in the conformation of the backbone although, quite naturally, the flexibility of the ethylene bridge was increased.

Superimposition of the β-turn mimetic from the lowest energy structure of 5.23 on an ideal type-II β-turn reveals that the backbone and side chain positions are fairly well mimicked with an rmsd of 0.43 Å (calculated from Cα of residue i to Cα of residue i+3). The only larger difference is found for the side chain of residue i+1 that differs approximately 60 degrees between 5.23 and an ideal type II β-turn, however since that position is occupied by a glycine in this case, this difference should only have a minor, if any, influence on the biological activity. When compared to the X-ray structure of Leu-enkephalin, the backbone rmsd values of the ensemble of structures calculated for 5.23 range from as low as 0.61 Å to 1.68 Å, with the flexibility of Leu5 being the main cause of the large fluctuation in the RMSD value (Figure 5.8). Despite the conformational resemblance between 5.23 and crystalline Leu-enkephalin, no binding to either the δ- or µ-receptor was detected (unpublished results).

**Figure 5.8.** Structural comparison between 5.23 (black) and the X-ray structure of Leu-enkephalin (grey). The conformer of 5.23 shown has a backbone rmsd of 0.61 Å compared to Leu-enkephalin.
5.4. Conclusions

The total absence of activity for the rigid \( \gamma \)-turn containing desmopressin analogues 5.3 and 5.4 compared to the more flexible analogue 5.5 that retains some activity suggests that desmopressin probably exhibits a different conformation than a \( \gamma \)-turn when bound to the \( V_2 \) receptor. The fact that analogue 5.5 has a 1000-fold lower activity than \([\text{Ala}^4]\)desmopressin, table 5.2, even though it displays a striking structural similarity to that found for desmopressin in aqueous solution, suggests that the amide bond between Phe3 and Ala4 modified in analogue 5.5 is important, either in the interaction with the receptor, or in stabilizing a \( \beta \)-turn in desmopressin, as proposed by Wang et al.[49]

In the case of Leu-enkephalin the incorporated \( \beta \)-turn mimetic was found to mimic the structure of a type-II \( \beta \)-turn fairly well, and the structure of 5.23 has a good resemblance to that of crystalline Leu-enkephalin but still analogue 5.22 showed no affinity for either the \( \mu \)- or \( \delta \)-opioid receptors. This observation together with other studies[72, 73] casts some doubts on the occurrence of a \( 4 \rightarrow 1 \) \( \beta \)-turn in the bioactive conformation of Leu-enkephalin, even though a \( \beta \)-turn is found in the X-ray structure of Leu-enkephalin[71] and has been suggested as the bioactive conformation by Bradbury et al[74]. Instead, a \( 5 \rightarrow 2 \) \( \beta \)-turn seems more promising.[72]

These results highlight how elusive the bioactive conformations can be when detailed knowledge of the receptor is unavailable and assumptions on the bioactive conformation has to be made solely on information acquired on the unbound peptide that in most cases are very flexible and thus can adopt many conformations. Still, the absence of biological activity is by no means a failure and provides useful information that can be used in future research. In addition, synthetic routes towards novel turn mimetics have been developed as well as a stereoselective synthesis of \( \text{Xxx}'\Psi[\text{CH}_2\text{O}]\text{Ala/Gly} \) isosteres.
6. Interaction Between Pilicides and Chaperones Involved in Pilius Assembly in Uropathogenic E. coli

6.1. Biological Background

In a time when multi-resistant bacterial strains is a growing problem, it is crucial to find new approaches that targets bacterial virulence. Adherence to the host tissue is one such process connected to bacterial virulence, critical to the pathogenicity of bacteria. The adherence process is in most cases mediated through a class of proteins called adhesins. In uropathogenic *Escherichia coli*, these adhesins are located at the end of pili, hair-like structures on the surface of the bacteria consisting of several different protein subunits. Inhibition of pilus biogenesis would effectively disarm the bacteria, making pilus assembly a good target for novel antibiotics. Type 1 and P pili are two classes of pili present in *E. coli* responsible for development of urinary tract infections, UTIs. Type 1 pili consists of a rod mainly composed of FimA subunits connected to a short tip fibrillum containing FimG and the adhesin FimH.[75] FimH has the ability to bind to mannose oligosaccharides present on the epithelial cells in the bladder,[76] a binding that is critical for the development of bladder infection, cystitis. P pili are constructed in a similar fashion with a rod consisting of repeating PapA subunits and a long, flexible tip fibrillum containing PapE, PapF and the adhesin PapG. PapG binds to Galα(1-4)Gal digalactoside receptors found in the kidney,[77] thereby initiating kidney infection, pyelonephritis. Although different in their subunit composition and their physical appearance, both type 1 and P pili
are assembled through the well-conserved chaperone-usher pathway, fig 6.1.

Figure 6.1. Overview of the chaperone-usher pathway responsible for the biogenesis of P pili in uropathogenic E. coli. The boomerang-shaped chaperone PapD (marked D) transports the subunit through the periplasm to the usher in the outer membrane. In absence of chaperone, the subunits activate the Cpx signalling pathway and are degraded.

The subunits all have an immunoglobuline(Ig)-like tertiary structure in their C-terminal domain but lack the seventh β-strand that would complete a canonical Ig-fold. This exposes a hydrophobic groove which renders them unstable in the absence of specialized chaperones in the periplasm. The periplasmic chaperones FimC and PapD, involved in the assembly of type 1 and P pili, respectively, are key components in pilus biogenesis since they are responsible for transport of the subunits through the periplasm to an usher in the outer membrane where the pilus is assembled. The structure of FimC has been solved with NMR
spectroscopy[78] and the structure of PapD with X-ray crystallography[79]. They exhibit a similar fold, containing two Ig-like domains with a cleft between them, forming an overall boomerang-like shape. The structures of the chaperones, when bound to a corresponding subunit or the C-terminal peptide from a subunit have been studied both with X-ray crystallography and NMR spectroscopy.[80-84] This revealed a common binding theme involving β-strand complementation between the G1 β-strand in the chaperone and the C-terminal β-strand in the subunit, thus completing the Ig-fold of the subunit. However, the β-strand interaction is parallel rather than antiparallel in contrast to the normal Ig-fold. In addition, a salt-bridge is formed between the C-terminal carboxylic acid of the subunit and Arg8 and Lys112 in the chaperone. These two conserved residues are present in both FimC and PapD and are located in the cleft between the two domains (Figure 6.2).

Figure 6.2. The NMR structure of periplasmic chaperone FimC.
When the chaperone-subunit complex reaches the usher, the chaperone is released through a mechanism called donor-strand exchange,[85] in which the parallel $\beta$-strand complementation of the chaperone G1 $\beta$-strand is exchanged for an antiparallel $\beta$-strand interaction with the N-terminal part of a subunit in the growing pilus. Blocking the chaperone would lead to aggregation and degradation of the expressed pilus subunits in the periplasm, thereby disrupting the whole pilus biogenesis.

Scheme 6.1. Pilicides selected for NMR studies. Pilicides 1-8 are based on the pyridone scaffold whereas pilicides 9 and 10 are based on tyrosine. Both types of pilicides are designed with the C-terminal peptide of the PapG subunit (11) as a starting point.

Based on the structural knowledge of the chaperone-subunit interactions, a number of pilus biogenesis inhibitors, pilicides, have been designed and synthesized.[86] The pilicides are designed to mimic the C-
terminal part of the PapG subunit with the aim of blocking the chaperone from forming the chaperone-subunit complex. Based on surface plasmon resonance data and hemagglutination studies, ten pilicides were selected for further NMR studies (cf., 1-10, Scheme 6.1).

### 6.2. NMR Studies of Chaperone-Pilicide Complexes

The aim of the NMR studies was to gather as much information as possible about the interaction between PapD, FimC, and the pilicides in order to gain knowledge about the depiliating effect they exhibit as well as aiding future design of pilicides. The binding affinity of the pilicides is of course one parameter that it is important to be able to estimate, at least to make an internal ranking of the pilicides if not in absolute numbers. We chose to do this with relaxation-edited experiments[87] which is an experiment that is fast and easy to conduct. The other and perhaps most important thing was to locate the binding site of the pilicides on the chaperone, since no information on the interaction surface can be derived from the biological assays performed. This was done by studying the chemical shift perturbations of the backbone amide groups of FimC in the presence of pilicide. Two additional ligand observed experiments, saturation transfer difference (STD) and transfer NOE experiments were also conducted but gave inconclusive results and were not pursued further.

#### 6.2.1. Relaxation-edited Experiments

The transverse relaxation times, $T_2$, for the individual species in a complex in fast exchange with their free state will appear as an average of the free and bound state:

$$R_{2\text{obs}} = p_f R_{2f} + p_b R_{2b}$$

(6.1)
Interaction Between Pilicides and Chaperones Involved in Pilus Assembly

Where $R_{2\text{obs}}$ is the observed transverse relaxation rate (the inverse of the $T_2$ relaxation time), $R_{2f}$ and $R_{2b}$ are the relaxation rates in the free and bound state, respectively, and $p_f$ and $p_b$ are the populations of the free and bound state. This can be used to detect binding of small ligands to their target protein in a simple one-dimensional experiment (Figure 6.3).

![Figure 6.3](image)

**Figure 6.3.** a) Pulse-sequence for a relaxation-edited experiment with WATERGATE solvent suppression. Narrow and wide bars corresponds to pulses with 90° and 180° flip angles, respectively. Grey rectangles are selective pulses. The length of the relaxation filter is determined by the value of $\Delta$ (usually 1-2 ms) and the number of loops, n. b) Resulting spectra using the pulse sequence in figure 6.2a for pilicide 2 with a relaxation filter of 200 ms. The upper spectrum is recorded in the absence of the chaperone PapD whereas one equivalent of PapD was added in the lower spectrum. Signals from 2 are marked with an asterisk. 1-Napthylacetic acid was used as a negative reference and its corresponding signals are unaffected by the chaperone.

A spin lock with a carefully chosen length after the 90° observation pulse acts as a relaxation filter, allowing the signals from the protein to disappear as a consequence of their fast $T_2$ relaxation. Small molecules, such as a ligand, will still be observable although with somewhat reduced signal intensity. A ligand bound to the protein will behave much as the protein during the time it is part of the complex. As a consequence, the $T_2$ relaxation time of a ligand will decrease as a function of binding strength, making it easy to detect binding by measurement of intensity of ligand
signals in absence and presence of protein, respectively.[87] Using this technique we could show that pilicides 1-7 and 9 displayed affinity for the chaperone PapD. Compounds 8 and 10 were not tested. The T₂ relaxation time for the pilicides in the presence of chaperone can be used to determine the dissociation constant, K_d. Individual R₂ relaxation rates was determined from a series of relaxation-edited spectra with the spin lock time varied between 0-4 s in the presence of one equivalent of PapD. Signal intensities were fitted to Equation 6.2.

\[ I = I_0 e^{-R_2t} \]  

(6.2)

I is the measured signal intensity at time t, I₀ is the signal intensity at t=0 and t is the length of the spin lock. Concentrations of the bound and free state of the pilicides were calculated using equation 6.1 and K_d could be calculated as

\[ K_d = \frac{(p_f [\text{pilicide}])^2}{(1-p_f) [\text{pilicide}]} \]  

(6.3)

where [pilicide] is the total concentration of the pilicide in the sample. Using this approach, all pilicides were found to bind to the chaperone PapD with dissociation constants ranging from 1 mM to 10 mM.

### 6.2.2. Assignment of FimC and Chemical Shift Mapping

Chemical shift mapping relies on the chemical shift perturbations of amide groups in the protein in the presence of a ligand, perturbations that when mapped onto the structure of the protein reveal the location of binding sites. An assignment of the backbone is thus a prerequisite for this experiment. We chose to work with FimC instead of PapD in these experiments since it is somewhat smaller than PapD and, more importantly, the assignment of FimC has been reported previously.[88] Both $^{15}$N-labelled and $^{13}$C/$^{15}$N-labelled FimC were
expressed from the *E. coli* strain BL21(DE3)pLys harbouring a modified pACA plasmid grown in minimal media containing $^{15}$NH$_4$Cl or $^{15}$NH$_4$Cl and $^{13}$C-glucose. Purification was performed using a DEAE cation exchange column followed by a eluation through a CM anion exchange column with a NaCl-gradient (0-0.5 M), yielding 8 mg/L labelled of both $^{15}$N-labelled and $^{13}$C/$^{15}$N-labelled FimC.

![Figure 6.4](image_url)

**Figure 6.4.** Contour plot showing $\omega_1$-$\omega_3$ strips from CBCACONH (left columns) and CBCANH (right columns) for residues Asn101 to Leu103. Horizontal lines show sequential connectivities between $^{13}$C$_{\alpha}$ (black in the CBCANH strips) and $^{13}$C$_{\beta}$ (grey in the CBCANH strips). The backbone $^{15}$N-shifts are shown at the top of each pair of strips.

As mentioned above, a complete assignment of FimC have been published earlier[88] but two mutations in the primary sequence of the FimC we purified, Glu→Val in position 18 and Thr→Ala in position 174, made it necessary to reassign the backbone. This was done with the triple-
resonance experiments CBCA(CO)NH[89, 90] and CBCANH[91] in conjunction with the earlier published assignment (Figure 6.4). The magnitude of the chemical shift perturbations, $\Delta$, was calculated as

$$\Delta = \sqrt{(^{1}H_f - ^{1}H_b)^2 + 0.2(^{15}N_f - ^{15}N_b)^2} \quad (6.4)$$

where $^{1}H_f$ and $^{15}N_f$ are the chemical shift of the amide proton in the reference spectrum containing only $^{15}N$-FimC and $^{1}H_b$ and $^{15}N_b$ are the corresponding chemical shift values in the presence of one equivalent of the different pilicides. The C-terminal peptide from the PapG subunit, 11, was used as a control substance in the chemical shift mapping experiment since its binding to the chaperone PapD has been well characterized, as discussed earlier. Since we used FimC in this study it could be argued that a peptide from a Fim subunit would serve better as a control but due to solubility issues with the C-terminal peptide from FimH, 11 was used. Furthermore, it has been shown that chaperones PapD and FimC are interchangeable in the pilus biogenesis of type 1 and P pili (Scott J. Hultgren, personal communication). The result from the control experiment was in excellent agreement with the crystallographic data. For the pilicides it became evident that a larger portion of the backbone is affected than expected from a single binding site located in the cleft between the two domains of the chaperone. The flexible loop connecting the $\beta$-strands F1 and G1 and the G1 $\beta$-strand itself, residues 92-115, are affected for all tested pilicides. Residues 150-153 and 189-191 located in or in close proximity to the cleft also exhibit changes in chemical shift in the presence of a majority of the studied pilicides. Taken together, these changes cover a large part of the chaperone-subunit interaction area. This led us to believe that direct binding of the pilicides to FimC is only partially responsible for the observed changes and that some other mechanism is also in operation. The chaperone PapD has been shown to be involved in a self-capping mechanism[92] where it forms a weakly bound homodimer that protects the subunit binding area when it’s not involved in subunit binding. The result that we observe could be explained if such a mechanism is present also for FimC, and the pilicide interfere with the FimC-FimC interaction. The line widths of the protein resonances are inconsistent with a stable dimer but $T_2$ relaxation times
derived from T1p data indicates that a small amount of dimer in fast exchange with the monomer indeed may exist since a general increase in T2 relaxation time is observed in the presence of peptide 11. In other words, FimC behaves as a smaller molecule although bound to the peptide. Dimerization, together with the fact that small, relatively hydrophobic molecules such as 1-10 probably interact mainly with the side-chains of the chaperone, makes it difficult to locate the actual binding site of the pilicides. The presence of secondary low-affinity binding sites could further complicate the picture. There are however differences in the pattern of residues affected by the different pilicides, differences that indicates that at least two binding sites are present on FimC (see below).

### 6.3. Principal Component Analysis of Chemical Shift Changes

A principal component analysis, PCA, of the Δ-values (equation 6.4) for each residue highlights the differences in chemical shift changes induced by the different pilicides. As can be seen from the score and loading plots from the resulting PCA-model (Figure 6.5), the pilicides can be divided into three groups containing 1, 3 and 4 (yellow), 2 and 5-7 (red) and 9 (blue), respectively (Figure 6.5a). The residues whose chemical shift changes are characteristic for pilicides 1, 3 and 4 are Ser29 (not affected for 3), as well as Phe55, Ser92 and Met93, all of which are located at the backside of the N-terminal domain of FimC near the F1-G1 loop. Pilicides 2 and 5-7 have an effect on residues Arg8, Ile111 and Lys112 located in cleft. Pilicide 9 induces relatively large chemical shift changes in a large part of the F1-G1 loop, including residues Lys95, Lys97, Asn101, Leu103, Gln104 and Ala 106 (Figure 6.5b).
Figure 6.5. a) Score plot showing the grouping of the pilicides. b) Loading plot.

Residues responsible for the grouping in the score-plot are marked with the corresponding colours.

Most of the residues are located close to the origo in the score-plot, indicating that any changes for these residues are independent of the pilicide added, supporting the dimerization mechanism discussed above. It cannot be ruled out however, that unspecific binding of the pilicides at the subunit-binding area of FimC could yield a similar result. Pilicide 10 was excluded from the PCA-model since it for some reason induced larger chemical shift changes than the other pilicides, thereby creating a biased model. From the observed chemical shift changes it could, however, be seen that 10 binds in a similar way as the structurally closely related pilicide 9.

From a competition experiment it became evident that pilicide 6 is able to interfere with binding of peptide 11 to FimC, a result consistent with the PCA-model. This was concluded after the observation that the chemical shift perturbations induced by peptide 11 on FimC was partially cancelled in presence of in ten equivalents of pilicide 6.
6.4. Conclusions and Discussion

Binding of the tested pilicides to the chaperone PapD was verified by relaxation-edited experiments. $K_{d}$-values were estimated from $T_2$ relaxation time measurements for pilicides 2, 3 and 5-7 in the presence of PapD and were found to range from 1 to 10 mM. Determination of the binding site of the pilicides on the surface of the chaperone FimC by the use of chemical shift mapping turned out to be more complex than expected. The large area affected in the protein upon addition of the different pilicides was bewildering at first and not consistent with the computer model of the pilicide-chaperone interaction.[86] This led us to suspect that the pilicides interfere with the weak association in FimC homodimers that are present under the conditions used for the chemical shift mapping experiments or that unspecific binding to the subunit-binding area of FimC occurs. The ability to form homodimers has been demonstrated for the closely related chaperone PapD.[92] Moreover, the interaction surface in the PapD dimer corresponds well with many of the chemical shift changes observed in our experiments for FimC. Presence of dimer formation for FimC was supported experimentally by measurements of $T_2$ relaxation times. The results are consistent with a small amount of weakly bound dimer that decreases upon addition of peptide 11. PCA of the chemical shift changes further supports the dimer disruption theory but also highlights specific differences in the interaction for the different pilicides with FimC. From the PCA-model result, it was concluded that two or possibly three different binding sites are present on the surface of FimC. One is located in the cleft between the two domains of the chaperone whereas the other is located in the N-terminal domain of the chaperone, including a part of the F1-G1 loop (possibly two adjacent binding sites). The ability to compete with the binding of peptide 11 was also experimentally verified by observation of diminished binding upon addition of ten equivalents of pilicide 6 to a 1:1 mixture of 11 and FimC. The pilicides binding in the cleft, as well as those binding in the loop, show a depiliation effect in a hemagglutination assay which indicates that two different mechanisms could cause depilation. One mechanism is inhibition of the chaperone by blocking the salt-bridge formation between the subunits and Arg8 and Lys112 located in the cleft. The other mechanism could be disruption of
chaperone-subunit complexes through binding of the pilicide to the binding site adjacent to the F1-G1 loop.
References

7. References


27. Kuszewski, J., A.M. Gronenborn, and G.M. Clore, *The impact of direct refinement against proton chemical shifts on protein*
References

References


References


Ett stort TACK till alla som på olika sätt bidragit till den här avhandlingen samt alla ni som gjort den här tiden till ett sant nöje:

**Ingmar Sethson**, min biträdeshandledare, som fick mig att inse att det här med NMR, det var något att satsa på. Din fingerfärdighet framför NMR instrumentet har varit en ovärdelig hjälp. Tack även för att du pushat för saker som jag initialt varit tveksam till (multivariat dataanalys?). Det enda du inte lyckats med är att övertyga mig om att Mac skulle vara bättre än PC.

**Jan Kihlberg**, min huvudhandledare, för att du gav mig den här chansen. Du har varit den som med jämna mellanrum har återfunnit den röda tråden och drivit projekten framåt när jag ibland har haft för många bollar i luften.


Rumsamrater och Labkamrater genom åren: **Tobias**, min kompanjon i källaren, **Jon**, min kompanjon när jag väl befördrades till plan 4 och **Mickael** som hållit mig sällskap på labbet.

**David**, för allt enkephalinsamarbete och glada tillrop, **Veronica**, för hjälp med proteinkrenning och sällskap vid NMR instrumentet, **Lotta** ”sämstboll?” för att du testade ”min” peptid, hela grupp **Fripppe** (pilicides Inc.),
Acknowledgements

Anna Linusson för värdefulla kommentarer till sista manuskriptet, Sharkmannen Tomas för alla duster på mariehemsängarna.

Organisk kemi, en härlig blandning av människor i vars lunchrum de mest osannolika samtalsämnen kan har behandlats, ibland till och med lite kemi.

Carina, som alltid har haft järmkoll på det administrativa.

Bert, för all service av NMR instrumenten, ditt engagemang i vindelävsloppet och alla hjälp med krånglande skrivare och datorer.

Jürgen, Göran och Anders för avdelningsöverskrivande NMR samarbete

Biokemisterna, för all hjälp när jag med viss skepticism stegade in bland inkubatorer och autoklaver. Särskilt tack till Viola för all assistans och till Micke som jag ”lånade” diverse pryler av.

Norsjöpajka: Kent, Jimmy, Johan, Tobias, Magnus och Lars (numera jänkare), i erat sällskap kan man verkligen släppa allt som har med jobb och kemi (förutom kanske ett visst ämne) att göra.

Alla trevliga människor i jongleringsföreningen tusen och en boll: Kalle, Elias, Anton, Anders och övriga. Finns det något bättre sätt att varva ner än att kasta saker vilt omkring sig?

Fooze, tack för alla trevliga stunder i cyberrymden de senaste åren.

Whiskyklubben Uisge Beatha, för alla livliga diskussioner kring dyrbara droppar (som dock ibland smakat både disktrasa och hyvelspån). Särskilt tack till Danne som fick mig att prova på discgolf.

Slutligen vill jag tacka mina föräldrar och mina bröder för allt stöd.