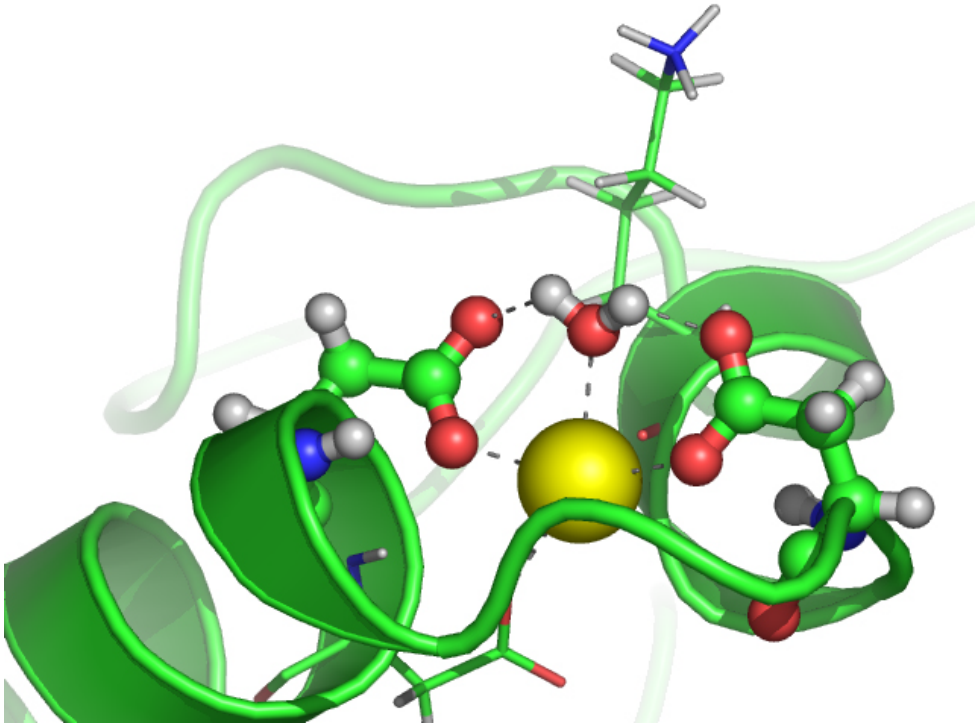
An NMR study of α-lactalbumin

pK_a-determinations and structural calculations



by

Jarl Underhaug

Submitted in partial fulfillment of the degree
candidatus scientarium

Department of chemistry
University of Bergen

Contents

Abstract

α-lactalbumin is a small protein (14 kDa) found in milk. It is one of the components of lactose synthase. Bacause of its metal ion binding properties, membrane binding properties, and its molten globule structure under certain conditions, it is a thoroughly studied protein.

The pH dependence of the interaction of calcium containing bovine α-lactalbumin with negatively charged membranes have revealed that interaction is dependent on electrostatic forces [1]. The protonation of D97 is believed to be of importance for the membrane interaction to occur [2].

To gain further insight into the initial step of the membrane interaction, the pK_a values of most of the aspartate, glutamate and histidine residues have been determined here. The pK_a values were determined by monitoring the pH dependent chemical shifts of protons close to the titratable groups. The pK_a values were also calculated theoretically, and compared with the experimental values.

There was in general a good agreement between the experimental and theoretical pK_a values. The exception was groups close to the flexible loop region in the C-terminal (L105-L110). The differences in this region are attributed to structural changes, which is not taken into account in the theoretical calculations.

D97, which is believed to be of importance for membrane interaction, has a low pK_a relative to pH 4.5 where the membrane interaction occurs.

An interesting observation is that many of the carboxyl groups with pK_a close to 4.5 are located in the same region of space. At pH 4.5 there will thus be a large change in the dipole moment of the protein. This change might orient the protein in the correct position for membrane interaction to occur.

Another interesting observation is that of the three calcium coordinating carboxyl groups, only D87 may be protonated when the calcium is bounded to the protein. The protonation of one more group will lead to loss of calcium.

As there does not yet exist an NMR structure of any α-lactalbumins, a structural calculation will be attempted here. Because an isotopically labelled protein was unavailable, the experiments were limited to homonuclear ^{1}H-^{1}H 2D experiments (COSY, TOCSY and NOESY), and ^{1}H-^{15}N HSQC at natural abundance. In addition to NOE restraints, the structural calculations also included torsional angle restraints derived from CSI, and $^3J_{\mathrm{H}^\alpha\mathrm{H}^\mathrm{N}}$-coupling constants extracted using ACME. The structural calculation was performed with the software package ARIA and CNS.

The calculated structure had a relatively low level of detail. The average RMSD was 1.6 Å for all backbone atoms, 1.1 Å for backbone atoms excluding the terminals, and 0.4 Å for the highly structured H3 α-helix. An NMR structure with average RMSD of 1.0 Å is regarded as a good structure, and a structure with an RMSD of 0.5 Å is regarded as a high quality structure. The main reason for this relatively low quality is the high flexibility of α-lactalbumin.

Acknowledgement

Let me begin by expressing my sincere gratitude to my supervisors, Nils Åge Frøystein and Aurora Martínez, for their guidance during the work with this thesis. Thank you, Nils Åge, this strange thing called nuclear magnetic resonance is finally starting to make sense. Your knowledge about NMR and computers have been vital for this project.

I'd like to thank Aurora for tipping me about the International School of Structural Biology and Magnetic Resonance. I had two great weeks at the summer school in Erice, Sicily, during the summer of 2003. It very was inspiring to meet other students, from all over the world, working within the same area as me. I made a lot of new friends there. A once in a lifetime experience! And of course a great thanks to Nils Åge too, it was after all he who gave me the financial support.

A great thanks to Beatriz Ibarra-Molero and José M. Sánchez-Ruiz at Universidad de Granada, Spain, for performing the theoretical pK_a calculations.

Also a thanks to the other members of The Biophysical Chemistry group too, especially Tormod Skauge for his help when I've had problems with my computer.

And not to forget all my friends, both at and outside of the department. I really appreciate your friendship, and I'm grateful for all the support you've given me. Last, but not least, a great thanks to my family.

And finally....
Thanks too everyone! I've had five great years at the university, or was it six years

Bergen, April 16th 2004

Jarl Underhaug

Abbreviations and symbols

Abbreviations

α-LA	α-LActalbumin
ACME	Amplitude-Constrained Multiplet Evaluation
ADR	Ambiguous Distant Restraint
ARIA	Ambiguous Restraints for Iterative Assignment
AUTOPSY	AUTOmated Peak Picking for NMR Spectroscopy
BASHD-TOCSY	BAnd Selective Homonuclear TOCSY
BLA	Bovine α-LActalbumin
CNS	Crystallography & NMR System
COSY	COrrelation SpectroscopY
CSA	Chemical Shift Anisotropy
CSI	Chemical Shift Index
CW	Continuous Wave
DSS	2,2-DimethylSilapentane-5-Sulfonic acid
DQE	Double Quantum Echo
DQF	Double Quantum Filter
EDTA	EthyleneDiamineTetraAcetate
FFT	Fast Fourier Transform
FHSQC	Fast Hetronuclear Single-Quantum Coherence
FID	Free Induction Decay
FT	Fourier Transform
GlcNAc	N-AcetylGlucosamine
GT	GalactosylTransferase
HSQC	Hetronuclear Single-Quantum Coherence
INEPT	Insensitive Nuclei Enhanced by Polarisation Transfer
MG	Molten Globule
MRI	Magnetic Resonance Imaging
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Enhancement SpectroscopY
RF	Radio Frequency
RMD	Restrained Molecular Dynamics
RMSD	Root Mean Square Deviation
TMS	TetraMethylSilane
TOCSY	TOtal Correlation SpectroscopY
UDP	Uridine DiphosPhate
WATERGATE	WATER suppression by GrAdient-Tailored Excitation

Symbols

a_{ab}	signal amplitude
b_{IS}	dipole-dipole coupling constant
$\mathbf{B_0}$	static magnetic field
$\mathbf{B_1}$	oscillating magnetic field
d_{IS}	secular dipole-dipole coupling constant
d_{ab}	distance detected by NOE, e.g $d_{\alpha\mathrm{N}}$
δ	chemical shift
E	energy
γ	gyromagnetic ratio
$\hbar$	Planck constant divided by 2π
η	viscosity
${}^nJ_{ab}$	*n*-bond coupling constant
$J(\omega)$	spectral density
k_B	Boltzmann constant
m_I	magnetic quantum number
μ	magnetic moment
μ_0	permeability in vacuum
ν	frequency in Hz
ω	frequency in rad/s
P	power
r_a	radius
r_{ab}	distance
R_{cross}	cross relaxation rate constant
σ	chemical shilding constant
t	time
t_1	incremental delay
T	temperatur
T_1	spin-lattice relaxation time constant
T_2	spin-spin relaxation time constant
τ_c	correlation time
τ_m	mixing time
τ_P	pulse length
V	volume
W	transition probability

Chapter 1

Introduction

1.1 Proteins

Proteins are one of the most important classes of biomolecules. They have a versatile role, ranging from building structures to catalyzing biochemical reactions. Together with DNA proteins form the basis for life as we know it.

There is a close relationship between proteins and DNA. The genetic information stored in the DNA codes for all the proteins needed to sustain life. DNA is also dependent on proteins for packing, replication, transcriptions and repair.

The majority of proteins in animals are found as structural proteins. This is fibrous proteins found outside of the cells, in the extracellular matrix. Tissues such as skin, bones, tendons, ligaments contain large portions of the protein collagen. About 20% of the muscle tissue consists of proteins, and their function depend on protein-protein interactions.

One important property of proteins is their ability to bind specific ligands. This is made use of in hormonal control, in the nerve system and in the immune system. Proteins in the cell membranes may trigger reactions inside the cells when exposed to the proper ligand on the outside.

Another class of proteins are the enzymes. The enzymes catalyze the biochemical reactions inside living organisms so that they are fast enough to sustain life. The enzymes have one or more sites where the substrate(s) bind. One purpose of these sites is to align the substrate molecules in the most favorable relative orientation for the reaction to occur. The substrates also bind more tightly in the transition state, thereby lowering the energy of the transition state and increasing the rate of the reaction. There are millions of different enzymes, each one catalyzes one specific reaction.

The protein studied in this thesis, α-lactalbumin, is a modulator. It changes the activity of another protein, galactosyltransferase.

1.1.1 Amino acids

The proteins are large polymers whose building blocks are α-amino acids. Amino acids are carboxylic acids which also carry an amino group. They are α because the amino group is on the carbon α to the carboxyl carbon. With the exception of glycine they all have a chiral center. One interesting point is that all amino acids found in nature are L-α-amino acids. This implies an S configuration for all amino acids except for cysteine.

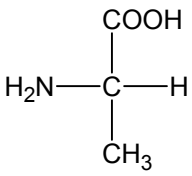


Figure 1.1: Fischer projection of a L-α-amino acid (alanine)

Another interesting property of the amino acids is the fact that they have both an amino and a carboxyl group. This makes them zwitterions in neutral solutions. In

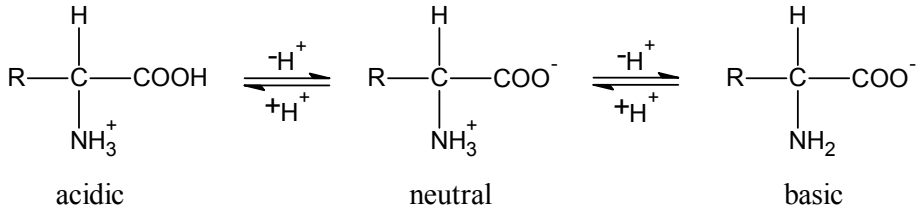


Figure 1.2: The ionization of amino acids at different pH-values

acid solutions they are positively charged, while they are negatively charged in basic solutions. In addition to the amino and carboxyl group the amino acids may also have functional groups on the side chains. More than 100 different amino acids have been found in nature, but only 20 of them are commonly found in proteins. In addition to their ordinary names, the 20 common amino acids are also given a one- and three-letter code (see Table A.1.10).

Table 1.1: One- and three-letter codes for amino acids

Amino acid	One letter	Three letter
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartate	D	Asp
Cysteine	C	Cys
Glutamine	Q	Gln
Glutamate	E	Glu
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

1.1.1.1 Hydrophobic amino acids

Amino acids are often classified by the properties of their side chains. A coarse classification is to classify them as either hydrophobic or hydrophilic amino acids. The hydrophobic amino acids can then be divided into aliphatic and aromatic. Alanine, valine, leucine and isoleucine belong to the aliphatic amino acids. Their most important

property is that they interact poorly with water. They do, however, interact with each other and other non-polar molecules. This is one of the main factors in stabilizing the protein structure. Methionine may also be classified as aliphatic. It is special because it has a sulphur atom between the methylenes and the methyl group. The aromatic hydrophobic amino acids are tryptophan, tyrosine and phenylalanine. Like the aliphatic amino acids they interact with other hydrophobic side chains and prefer to be buried in the hydrophobic cores of the protein.

In addition to the already mentioned hydrophobic amino acids, there are a few special ones. Glycine is the most simple. It has no side chain, except for a hydrogen atom. Proline has an aliphatic side chain, but the chain is bounded back to the nitrogen on the backbone. This really makes proline an imino acid, and not an amino acid. The ring imposes a constraint on the C-N bond. Because of this, proline has a significant effect on the backbone conformation. The last of the hydrophobic amino acids, cysteine, has a very important property regarding protein structure: the ability to form covalent bonds with another cysteine residue (see Figure 1.3). These covalent bonds stabilize the tertiary structure (see subsection 1.1.4) of proteins.

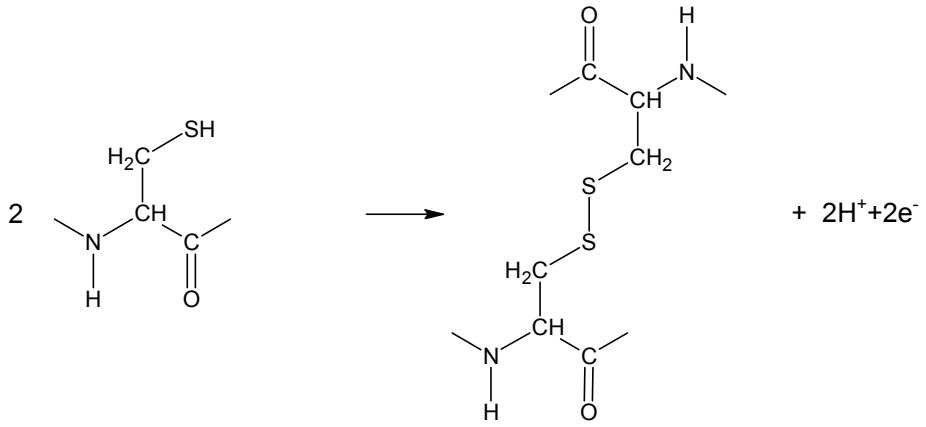


Figure 1.3: The formation of a disulfide bridge between two different cysteine residues. This stabilizes the tertiary structure of proteins.

1.1.1.2 Hydrophilic amino acids

Having a pK_a near physiological pH, histidine is a flexible basic amino acid. It may either be buried within the protein or exposed to the solvent. In addition, the flexibility makes it ideal for functional centers. The other basic amino acids, lysine and arginine, are considered amphipathic because of their long hydrophobic chain with a positive charge on the end. This property is the reason why they may both be either on the outside of proteins or partly buried within them. Together with the acidic amino acids aspartate and glutamate[1], lysine and arginine are frequently involved in ionic bridges. These bridges are important for protein stability. Having a pK_a of around 4, aspartate and glutamate are charged at physiological pH. Being charged, aspartate and glutamate prefer to be on the surface of the protein, exposed to an aqueous environment. The charge is also the reason for their ability to bind certain metal ions.

[1] aspartate and glutamate are sometimes referred to as aspartic and glutamic acid

Asparagine and glutamine are the amide forms of the acidic amino acids. Being polar, they prefer to be on the surface of the protein. They may each form 4 different hydrogen bonds, and are therefore important in stabilizing the protein structure. Another group of polar amino acids are the hydroxyl residues serine and threonine. They are not as polar as asparagine and glutamine. Therefore they may also be found on the inside of the protein.

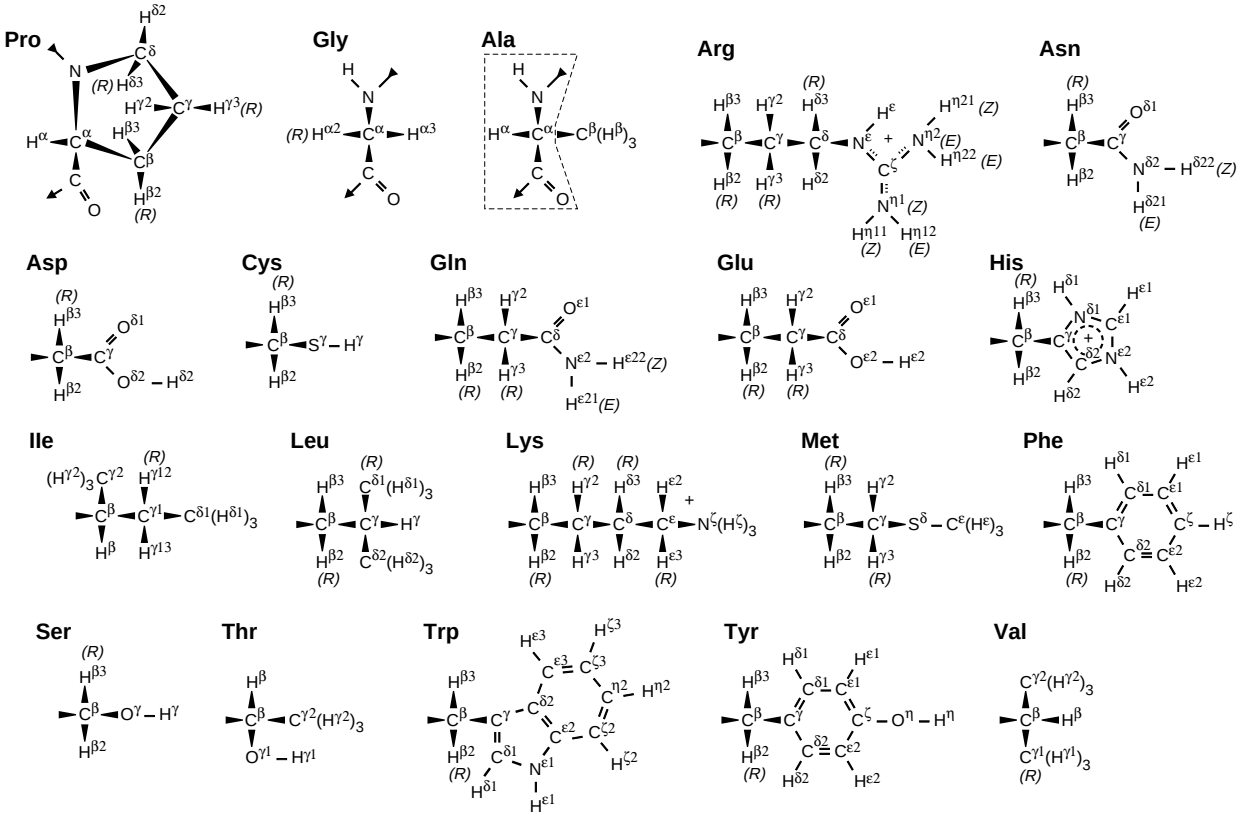


Figure 1.4: The structure of the 20 common amino acid side chains. The nomenclature recommended by IUPAC [3] is also shown. H^N is the recommended nomenclature of the backbone amide proton.

1.1.2 Primary structure

Two amino acids may bond together by the formation of a peptide bond. It is formed between the amino group of one amino acid and the carboxyl group of another amino acid (see Figure 1.5). In this process one molecule of water is lost. When many amino acids are connected through peptide bonds, we get a polypeptide. To be called a protein a polypeptide is usually required to consist of more than 50 amino acids, be naturally occurring and have a definite three-dimensional structure.

The peptide bond has an important property regarding the rigidity of the polypeptide chain and consequently the folding of the protein: it has a partial double bond character due to resonance effects. The peptide bond length is 1.33 Å, shorter than

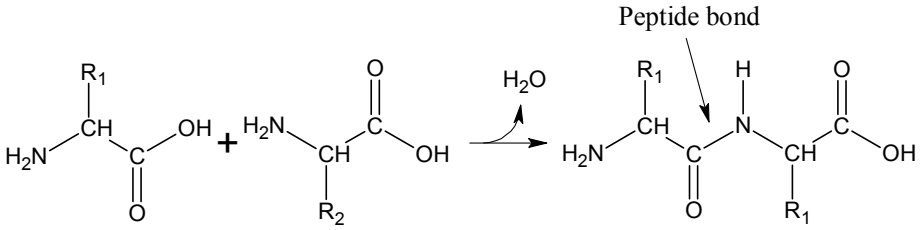


Figure 1.5: Two amino acids are linked together by a peptide bond

the length of a normal C-N bond. This double bond character of the peptide bond has been calculated to 40%. Because of the double bond character, the six atoms in Figure 1.6 are all in virtually the same plane. The trans form shown in the figure is the most common in proteins.

The primary structure of a protein is simply the sequence of amino acids. By convention the sequence is written from the N-terminal to the C-terminal end of the polypeptide chain. The N-terminal amino acid is defined as number 1. Using the single-letter code, the primary structure of human myoglobin would be:

```
GLSDGEWQLV LNVWGKVEAD IPGHGQEVLI RLFKGHPET LEKFDKFKHL
KSEDEMKASE DLKKHGATVL TALGGILKKK GHHEAEIKP LAQSHATKHK
IPVKYLEFIS ECIIQVLQSK HPGDFGADAQ GAMNKALEL FRKDMASNYK
ELGFQG
```

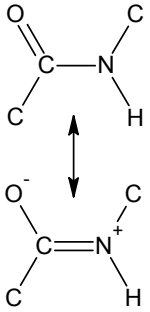


Figure 1.6: Resonance of the peptide bond

1.1.3 Secondary structure

The secondary structure of a protein is concerned with the local three-dimensional arrangement of the polypeptide backbone. The packing of the side chains is not an issue here. Because the atoms in the peptide bond are virtually in the same plane, there are only two torsion angles for each amino acid that determine the structure of the backbone. These two angles are designated ϕ and ψ (see Figure 1.7). The torsion angle of the peptide bond is called ω.

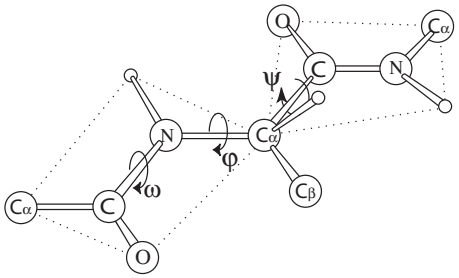


Figure 1.7: Definition of torsion angles in the polypeptide backbone. The dotted lines indicate the peptide planes.

The secondary structure is stabilized by hydrogen bonds that are formed between different amino acids. Each amino acid has two potential sites for hydrogen bonding on the backbone: C=O and N-H.

There are two main secondary structures: the α-helix and the β-sheets. These structural units are connected by β-turns or random coils.

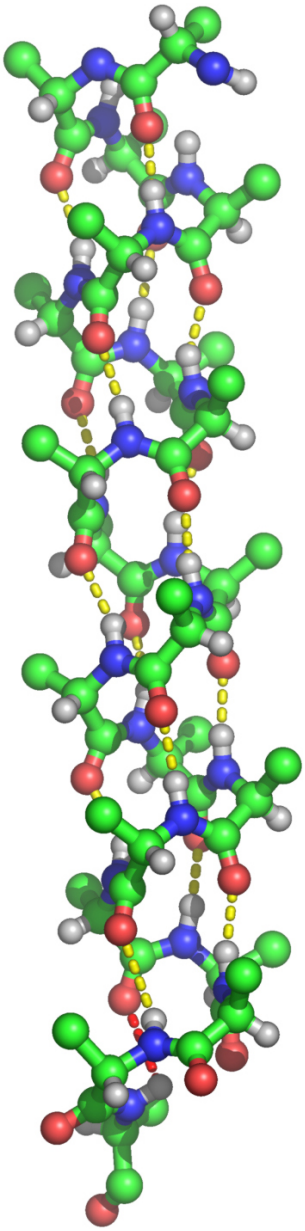


Figure 1.8: Right-handed α-helix. Dotted lines indicate hydrogen bonds.

1.1.3.1 α-helix

The most frequently found secondary structure is the α-helix. In the α-helix the polypeptide backbone is coiled about a central axis. For the L-amino acids the right handed helix is the most stable. The helix is stabilized by hydrogen bonds parallel to the axis, between C=O on amino acid *i* and N-H on amino acid *i+4*. To enable such hydrogen bonding, it takes 3.6 amino acids to complete one turn of the helix. The ϕ and ψ angles are virtually the same for all residues, -57° and -47° respectively. However, natural proteins have been found to deviate slightly from these values. The side chains are pointing outwards from the helix, directed slightly towards the amino end of the helix.

There also exist two variations of the α-helix. These helices are stabilized by hydrogen bonds between amino acid *i* and *i*+3 and between *i* and *i*+5. These variations of the α-helix are designated 3_{10}-helix and π-helix, respectively. The tighter and looser packing of these helices makes them unfavorable compared with the α-helix. The π-helix is an extremely rare secondary structural element in proteins.

1.1.3.2 β-sheets

The other major secondary structures are the β-sheets. There are two variants of the β-sheet, the parallel and the antiparallel. The basic unit for both these structures is the β-strand. In this structure the backbone is almost fully extended. The β-strand may also be viewed as a helix with 2.0 residues per turn. In this configuration the backbone has no hydrogen bonds, which makes the β-strand unstable on its own. To be stabilized the β-strand has to make hydrogen bonds with another β-strand, either in a parallel or antiparallel fashion. The peptide planes form a hill-and-valley pattern, but on a larger scale the β-strands are almost planar. The side chains are alternately pointing above and below the plane. The ϕ and ψ angles are -119° and 113° for the parallel β-sheet and -139° and 135° for the antiparallel.

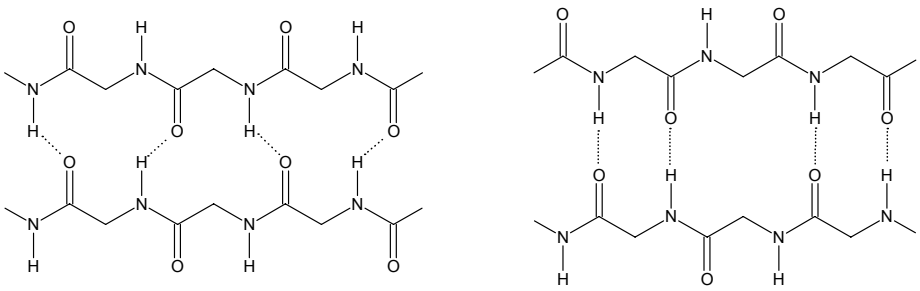


Figure 1.9: The β-sheets: parallel to the left and antiparallel to the right. The dotted lines indicate hydrogen bonds. Side chains and H^{α} are not shown.

1.1.3.3 Other structures

β-turns are found connecting the sheets in antiparallel β-sheets or connecting two α-helices. The β-turns are related to the 3_{10}-helix in the sense that there is hydrogen bond between the amino acids *i* and *i*+3. There are four different types of β-turns, type I and II, and their mirror images I' and II'. The number of amino acids participating in a β-turn varies, but four is a typical number.

Other less common regular structures include poly(Pro) and poly(Gly). Two variations exist of both these structures.

1.1.4 Tertiary structure

The various secondary structural elements are arranged into compact globular structures called domains. Larger proteins may consist of many domains linked together by a single polypeptide chain. The arrangement of the secondary structures into compact spheres is defined as the tertiary structure. While the secondary structures are stabilized by hydrogen bonds in the backbone, the tertiary structure is stabilized mostly by interactions between the side chains far from each other in the primary sequence, but close in space.

The side chains of many of the amino acids are capable of forming hydrogen bonds with each other. Glutamine and asparagine are important in this context. They are both proton donors and acceptors. They may also form "long-distance" hydrogen bonds with the backbone. Other amino acids capable of forming hydrogen bonds are aspartate and glutamate, which are proton acceptors, and lysine and arginine, which are donors. Serine, threonine and tyrosine may be both acceptors and donors. In addition to hydrogen bonds, the charged amino acids may also form true ionic bonds.

The strongest interaction stabilizing the tertiary structure is the disulphide bridge. A disulphide bridge is a covalent bond between two cysteine residues in different parts of the polypeptide chains (see Figure 1.3).

The last major stabilizing factor is the hydrophobic interaction. This "force" results from the unfavorable constraints placed on water as it packs around a non-polar hydrocarbon chain. The water molecules are constrained, hence, the entropy of the system decreases. Consequently, the hydrophobic chains will pack together and stabilize the protein structure.

In certain proteins the tertiary structures aggregate and form multimeric units. This is called the quaternary structure.

1.2 α-lactalbumin

α-lactalbumin (α-LA) is a small protein (14 kDa) found in milk. It is one of the components of lactose synthase. Because of its metal ion binding properties and its molten globule structure under certain conditions it is a thoroughly studied protein. Recently it has also been discovered that α-LA has bactericidal and anti-tumor properties.

1.2.1 Structure

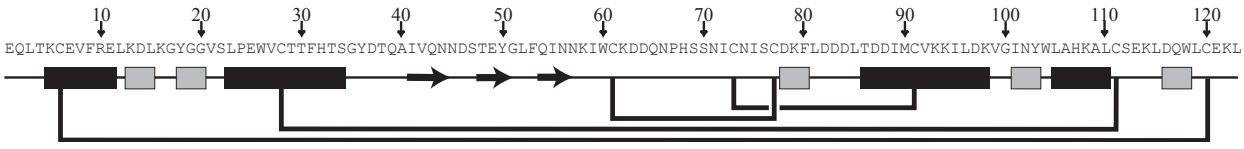


Figure 1.10: The bovine α-lactalbumin sequence with disulfide bridges and secondary structures (black boxes=α-helices, gray boxes=3_{10}-helices, and arrows=β-sheets)

Most α-LAs consist of 123 amino acids, that includes the bovine α-LA (BLA) which is used in this thesis. The sequence is homologous with that of the c-type lysozymes [7]. The tertiary structure is also very similar [8].

α-LA consists of two domains, a large α-helical domain and a smaller β-sheet domain, separated by a cleft. In the α-helical domain there are four α-helices (H1:5-11, H2:23-34, H3:86-98 and H4:105-110) and four small 3_{10} helices (h1a:13-15, h1b:18-20, h3:101-103 and h4:116-118). The β-sheet domain consists of a three-stranded antiparallel β-sheet (S1:41-44, S2:47-50 and S3:55-56), one small 3_{10} helix (h2:77-80) and a large unstructured region [4, 9, 10] (see Figure 1.11). In between the two domains the strong calcium-binding site is located (see next subsection).

Four disulfide bridges stabilize the structure, three internal in the domains (6-120, 61-77 and 28-111) and one connecting the domains (73-91).

1.2.2 Metal ion binding

α-LA has several different metal ion binding sites. The primary site is a 7-coordinated Ca^{2+}-binding site with an association constant in the order of 10^7 M^{-1} ($1.3 \cdot 10^7$ M^{-1} for BLA at pH 7.53 and $T = 40$ °C [11]). When calcium is bound to the protein it is referred to as "holo", without calcium it is referred to as "apo". The calcium ion is coordinated by five amino acids, the carboxyl oxygens of the residues D82, D87 and D88 and the backbone carbonyl oxygens of residue K79 and D84 (see Figure 1.12). Two water molecules are also participating in the coordination. The oxygen atoms form a pentagonal bipyramid [4, 10]. The structure is similar to the EF hand. At low pH values there is a competition between the calcium ion and the carboxyl protons:

$$\mathrm{P} + \mathrm{Ca}^{2+} \rightleftharpoons \mathrm{PCa}$$

$$\mathrm{P} + n\mathrm{H}^{+} \rightleftharpoons \mathrm{PH}_n$$

The pK_a of the three carboxyl groups have been reported to be 5.0 ± 0.1 [12, 13]. The binding of a Ca^{2+} ion to this site significantly increases the stability of the protein. Mutating the D87 residue participating in the calcium binding has been shown to alter the calcium-binding affinity and the protein stability [14]. α-LA has a secondary calcium binding site, 7.9 Å from the primary site [15].

α-LA also has one strong and several weak zinc binding sites. In human α-LA the only site verified has been found in the cleft region, bonded to E49 [16]. Site-directed mutagenesis showed that E49 was not involved in the binding BLA, but E1 was [17].

In addition to calcium and zinc, α-LA has also been shown to bind several other metal ions [13, 18]

1.2.3 Primary physiological function

The primary physiological function of α-LA is to regulate the production of lactose in the mammary secretory cells, through modifying the function of β1,4 galactosyltransferase (GT). In the absence of α-LA GT catalyzes the transfer of galactose from UDP-galactose to glycoproteins containing *N*-acetylglucosamine (GlcNAc). The binding of α-LA to GT modifies its substrate specificity from GlcNAc to glucose. Galactose will now be transferred to glucose, thus lactose will be produced:

$$\text{UDP-galactose} + \text{glucose} \rightarrow \text{lactose} + \text{UDP}$$

For optimal function lactose synthase, the complex of α-LA and GT, requires Mn^{2+} ions. The activity is also affected by Zn^{2+} ions, but the role of calcium is not known.

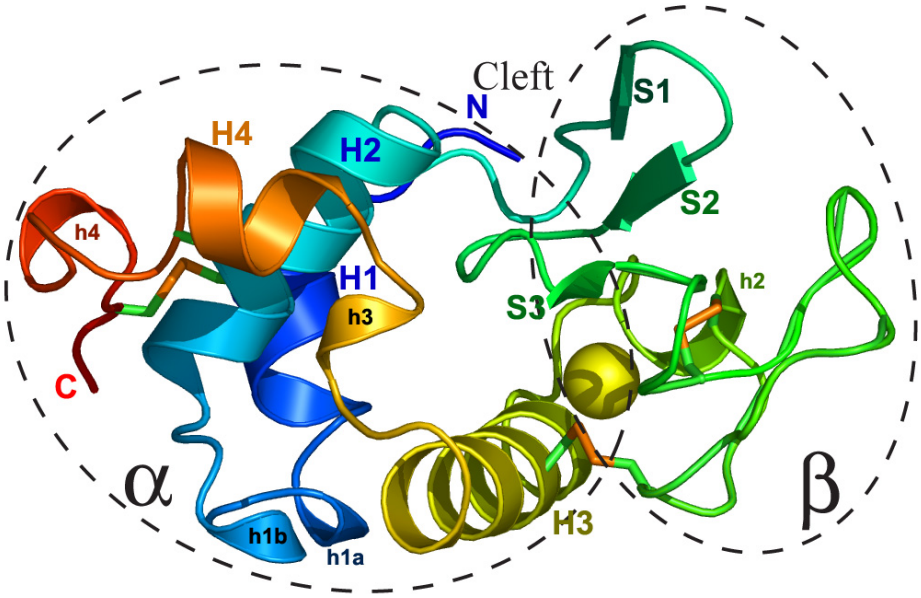


Figure 1.11: Cartoon representation of the x-ray structure of bovine α-lactalbumin [4] (PDB ID *1F6S* in the RCSB Protein Data Bank [5]). The α-helices are denoted by "H", 3_{10}-helices by "h" and β-sheets by "S". The α-helical and β-sheet domain are indicated by dotted lines. Disulfide bridges and the Ca^{2+} ion is also shown. This figure, and the rest of the protein graphics are created using the PyMOL software [6].

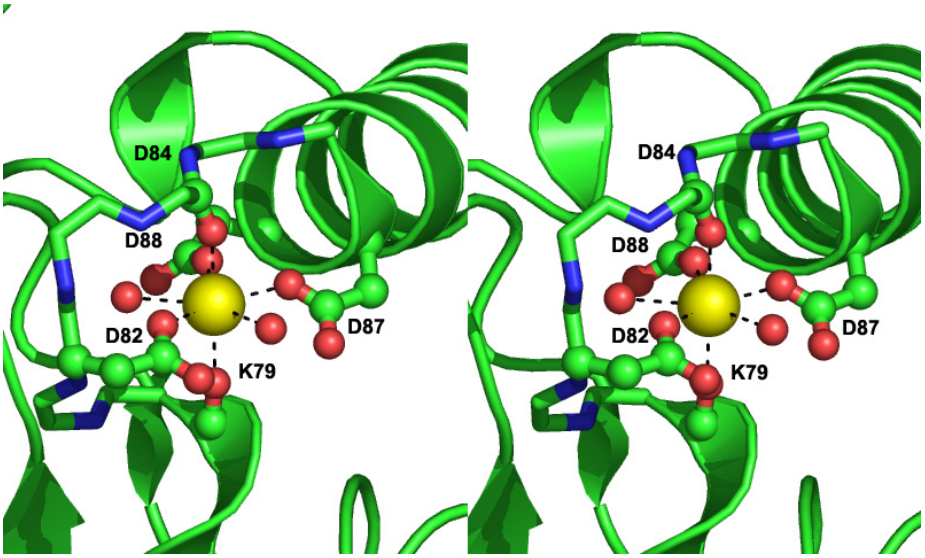


Figure 1.12: Stereoview (cross-eye) of the primary calcium binding site of BLA.

1.2.4 Other physiological functions

In addition to its primary function, α-LA also has bactericidal activities. Proteolytic digestion of BLA by trypsin and chymotrypsin yield three polypeptide fragments with antibacterial properties, primarily against Gram-positive bacteria [19]. Håkansson et al. [20] have described a molten globule-like folding variant of human α-LA, stabilized by oleic acid, which has bactericidal activity against *Streptococcus pneumoniae*. The same folding variant has also been shown to induce apoptosis in tumor cells and immature cells [21]. The conditions under which this complex forms is similar to the environment in the stomach of a nursing child. It is proposed that this conformation may lower the incidence of cancer in breast-fed children. It is believed that HAMLET induces apoptosis through interaction with histones and chromatin in tumor cell nuclei [22].

1.2.5 Molten globule

Some proteins have been shown to, under certain conditions, have a conformation that is neither fully folded nor unfolded. α-LA is one of the proteins that has such a conformation, and it is one of the best characterized. The state may be induced by temperature (T state), low pH (A state), mild denaturants (P state) and removal of the calcium ion (apo-state) [23]. The temperature, pH and amount of denaturants needed to induce the molten globule conformation is critically dependent on the metal ion concentration.

In the molten globule the secondary structures are retained, but the tertiary structure is slowly fluctuating [23]. The conformation is highly hydrated [24], which is reflected by an increased radius [25]. There is also evidence of an acid induced pre-molten globule state, an intermediate between the native structure and the A state [26]. NMR spectra of the molten globule structure have a limited chemical shift dispersion, but not as limited as that expected for a random coil conformation [27]. The study of the molten globule is interesting because it is believed to be a folding intermediate.

1.2.6 Interactions with membrane systems

It has be shown that α-LA can reversibly associate with membrane systems. The understanding of this association is important because it may increase the understanding of the physiological function of α-LA where it interacts with GT on the membrane surfaces in the Golgi lumen.

The initial step of the association with a negatively charged membrane is aided by electrostatic forces. Holo-BLA interacts with membranes at pH 4.5 and below because of dominating electrostatic repulsions at higher pH. The following step is believed to be driven by hydrophobic interactions. Apo-BLA also associates with membranes at higher pH because of exposure of apolar patches at the protein surface [1].

Membrane bound BLA has an increased α-helical content compared with the native structure. The tertiary structure is loose, similar to that of the MG state [28].

The protonation of the E11 and D97 residues are believed to be critical for the membrane association. Electrostatic attraction between the negative membrane and the positive residues on helix H1 and H3 will bring the protein closer to the membrane. As the protein approaches the membrane helix H1 and H3 twist exposing the hydrophobic residues which then interact with the membrane [2]. The increased helical content of BLA bound to the membrane may be related to the involvement of helix H3 in the membrane binding [29].

1.3 The aims of the work

Holo-BLA has been shown to interact with negatively charged membranes through electrostatic forces. As the interaction occurs around pH 4.5 the protonation of aspartate and glutamate is believed to be essential. To gain further insight into the membrane interaction of BLA, one of the goals of this work was to determine the pK_a values of the aspartic and glutamic acids.

In this context NMR is an indispensable tool. The pK_a values of all residues can be determined at once through the study of the pH dependency of chemical shifts of the residues. There are several types of experiments which can be used, both homonuclear and heteronuclear experiments. The choice of experiment was influenced by two factors:

- BLA is available as lyophilized powder from Sigma-Aldrich. This is prepared from raw, unpasteurized milk, and is obviously nonlabelled.
- The plasmid to express BLA isotopically labelled with ^{13}C and ^{15}N was not available when this project was started.

Based on the necessity to work with the BLA from Sigma-Aldrich, the choice of experiments is limited to homonuclear experiments. The most obvious choice then is a TOCSY-type experiment.

Despite a number of NMR investigations of α-lactalbumin by some groups [30–32], no solution structure of this protein has been deposited in the RCSB Protein Data Bank [5]. Therefore, another goal of this work was to determine the structure of BLA in solution by NMR. The experiments used to determine the structure were limited to homonuclear ^{1}H-^{1}H 2D experiments and heteronuclear ^{1}H-^{15}N correlation methods at natural abundance.

Chapter 2

Theory and methods

2.1 NMR-spectroscopy

In the beginning of the 20th century there were several discoveries leading up to the discovery of nuclear magnetic resonance (NMR). The Zeeman effect introduced magnetic fields as a part of spectroscopy, and the hyperfine structure of spectral lines led Pauli in 1924 to propose that atomic nuclei have an intrinsic angular moment. In 1933 the magnetic moment of the proton and deuteron in hydrogen molecules were determined using a molecular beam. Several different attempts were made to make use of this, and finally, in 1945, NMR was discovered simultaneously by the groups of Bloch [33, 34] and Purcell [35]. The later discovery of the chemical shift and the use of Fourier transform spectroscopy has, among other breakthroughs, increased the use of NMR. Today the use of multiple dimensions and increasingly higher magnetic fields has made NMR an indispensable analytical tool.

NMR has also found its use as a diagnostic method in medicine. The combination of NMR and magnetic field gradients has lead to new imaging tool, MRI.

2.1.1 Basic NMR theory

Many nuclei possess a nuclear spin, *I*, and thereby a magnetic moment, μ. When placed in a magnetic field $\mathbf{B}_0$, aligned along the z-axis, the energy of these nuclei is given by the equation:

$$\begin{aligned} E &= -\mu_z B_0 \\ &= -m_I \gamma \hbar B_0 \end{aligned} \tag{2.1}$$

where μ_z is the z-component of the magnetic moment, γ is the gyromagnetic ratio, $m_I = I, I-1, .., -I$ is the magnetic quantum number and $\hbar$ is the Planck constant divided by 2π. In a nucleus with spin $\frac{1}{2}$ there are two energy levels, $m = +\frac{1}{2}$ and $m = -\frac{1}{2}$, corresponding to the magnetic moment being aligned either along or against the magnetic field. These two energy levels are also referred to as α and β, where α is lowest energy level. The magnetic moments are not directly parallel/antiparallel with the magnetic field, but have an angle $\theta = 54°44'$ with the z-axis (see Figure 2.1). They will precess about the z-axis with the Lamor frequency (see equation 2.4). At

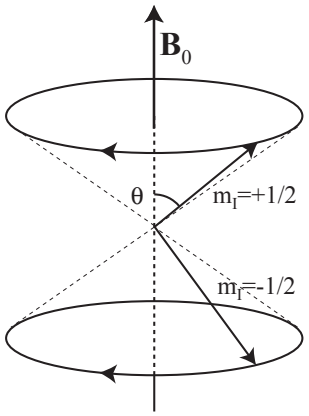


Figure 2.1: The two precessional cones for a spin-$\frac{1}{2}$ nucleus.

equilibrium they are randomly distributed around the two cones. Their distribution between the two energy levels is given by Boltzmann statistics:

$$\frac{N_\beta}{N_\alpha} = e^{-\frac{\Delta E}{k_B T}} \tag{2.2}$$

The energy difference between the to levels are:

$$\Delta E = \gamma \hbar B_0 \Rightarrow \quad \nu = -\frac{\gamma B_0}{2\pi} \tag{2.3}$$

$$or \quad \omega = -\gamma B_0 \tag{2.4}$$

This is called the resonance condition, and the frequency is called the Lamor frequency. Because the energy difference low, the populations of the two energy levels are almost equal. This is the reason for the low sensitivity of NMR. According to quantum mechanics, the only allowed transitions are those with $\Delta m = \pm 1$. Therefore, this is the energy that may be absorbed or emitted by all nuclei.

NMR would be of little use if all nuclei of one kind had exactly the same frequency. Luckily, the frequency is also dependent on the chemical environment. When placed in a magnetic field the electrons will start to circulate around the nucleus. This circulation will set up a new magnetic field in the opposite direction of $\mathbf{B}_0$. The nucleus will then experience a weaker field that is dependent on the electron density:

$$\nu = -\frac{\gamma B_{\text{eff}}}{2\pi} = -\frac{\gamma(1-\sigma)B_0}{2\pi} \tag{2.5}$$

σ is called the shielding constant. To be able to compare results from different spectrometers with different magnetic fields, a relative scale is used:

$$\delta = \frac{\nu_{\text{sample}} - \nu_{\text{ref}}}{\nu_{\text{ref}}} 10^6 \tag{2.6}$$

The δ-values are reported in ppm. For ^{1}H and ^{13}C tetramethylsilane (TMS) is used as a reference. By definition TMS has a δ-value of zero. One problem with TMS is that it is insoluble in water. If water is used as a solvent, another reference compound has to be used and the results converted to the TMS scale.

2.1.2 CW NMR method

The simplest method of recording a NMR spectrum is to use the continuous wave (CW) method. A transmitter coil perpendicular to the magnetic field constantly irradiates the sample with a frequency that is slowly varied to cover the entire spectrum. At the same time a receiver coil, perpendicular to both the transmitter coil and the magnetic field, detects the radio frequency (RF) radiation that has been absorbed and emitted by the sample. This method is both time consuming and has a low sensitivity. The experiment may also be performed by varying the magnetic field. Until 1966, when Ernst and Anderson developed the pulsed Fourier transform (FT) method [36], this was the only way of recording a spectrum.

2.1.3 FT NMR method

Compared with FT NMR, the CW method is very inefficient. While a CW spectrum takes several minutes to record, the same FT spectrum may be recorded in a few

seconds. In addition to the time factor, the ability to use low sensitivity nuclei and higher dimensions makes FT NMR the only method considered today. For a schematic overview of a modern FT NMR spectrum, see Figure 2.2.

When describing FT NMR experiments it is most convenient to work with the macroscopic magnetization vector, $\mathbf{M}_0$. This is the sum of the magnetizations of all the individual spins. At equilibrium the vector is aligned along the z-axis. Its magnitude is given by:

$$M_0 = (N_\alpha - N_\beta)\frac{1}{2}\gamma\hbar = \frac{N(\gamma\hbar)^2 B_0}{4k_B T} \tag{2.7}$$

This vector may be flipped around by RF pulses.

2.1.3.1 The RF pulse

In the pulsed NMR method a short RF pulse is used to excite all the nuclei of the same species simultaneously (see Figure 2.3). This pulse has a frequency ν. However, if the pulse is very short it contains, according to Heisenberg's principle of uncertainty[1], a continuous band of frequencies, not only the frequency ν (see Figure 2.4). The width of the frequency band is proportional to τ_P^{-1}, where τ_P is the length of the pulse. ν is set to the center of the spectrum. τ_P should be short enough to uniformly excite all nuclei of interest. Usually τ_P is in the range of a few μs.

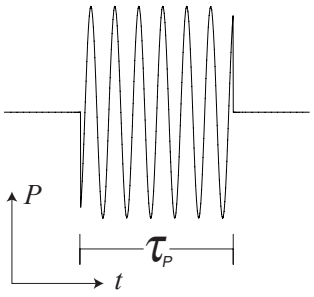


Figure 2.3: The RF pulse

The RF pulse is generated by an auxiliary oscillating field, $\mathbf{B}_1$, perpendicular to $\mathbf{B}_0$. This $\mathbf{B}_1$ field may be decomposed into two magnetic fields, $\mathbf{B}_{1r}$ and $\mathbf{B}_{1l}$, rotating in the xy-plane. They both rotate with the frequency ν, but in different directions. Only the magnetic field rotating in the same direction as the nuclear dipoles interacts with macroscopic magnetization vector to a considerable degree. The visualization of the vector model becomes much more simple if we introduce a rotating coordinate system, (x', y', z) (see Figure 2.5). This coordinate system is defined to rotate in the same direction as the dipoles and with frequency ν.

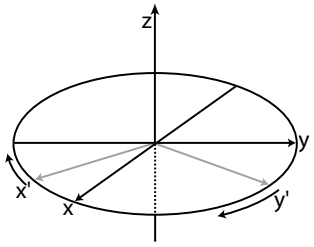


Figure 2.5: The rotating coordinatesystem

An x-pulse is defined to have its effective component of the magnetic field aligned along the x'-axis of the rotating coordinate system. The pulse will rotate $\mathbf{M}_0$ about the x'-axis. The angle through which it will be rotated, the flip angle, is given by the equation:

$$\Theta = 2\pi\gamma B_1 \tau_P \tag{2.8}$$

A 90° x-pulse will rotate $\mathbf{M}_0$ down to the y'-axis. Instead of being randomly distributed, the magnetizations of all the individual spins will now precess in phase, and the populations of the two energy levels will be equal.

2.1.3.2 Relaxation

After the magnetization has been disturbed by a pulse it will, over time, return to thermal equilibrium. This process is called relaxation. Relaxation may be divided in to different types: relaxation parallel to the external magnetic field and relaxation perpendicular to the field. The two different types are described mathematically by the Bloch equations:

$$\frac{dM_z}{dt} = -\frac{M_z - M_0}{T_1} \tag{2.9}$$

$$\frac{dM_{x'}}{dt} = -\frac{M_{x'}}{T_2} \quad \wedge \quad \frac{dM_{y'}}{dt} = -\frac{M_{y'}}{T_2} \tag{2.10}$$

[1] $\Delta E \Delta t \geq \frac{\hbar}{2}$

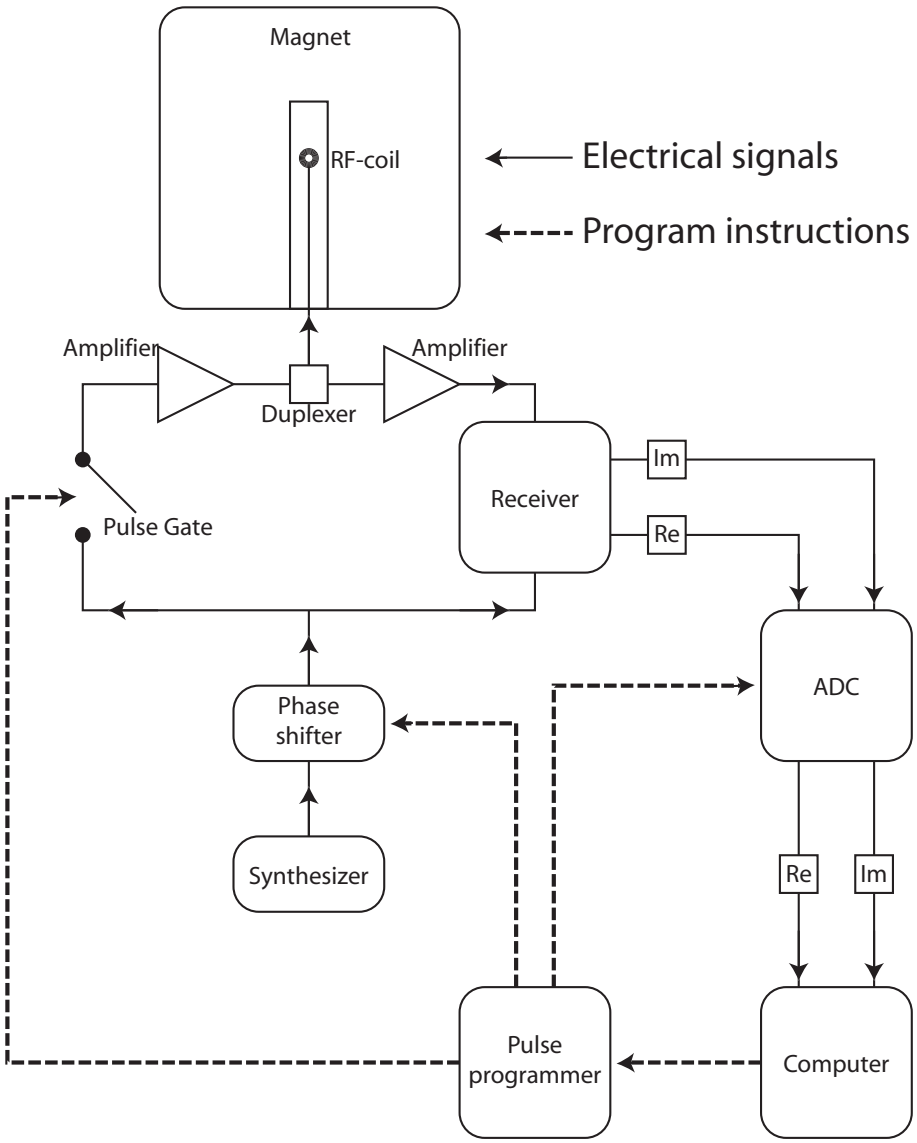


Figure 2.2: Schematic overview of a FT NMR spectrometer (adapted from Spin Dynamics [37]). The Synthesizer generates the RF wave. This wave is phase shifted according to the instructions in the pulse program. The pulse gate lets the wave through to the probe during the pulse period. The FID is directed to the receiver, where the high frequency signal is converted to two audio frequency signals by substraction of a reference frequency. These signals are digitalized by the AD converter and fed to the computer. The role of the duplexer is to direct the high power signal to the probe. The low power FID is directed to the receiver.

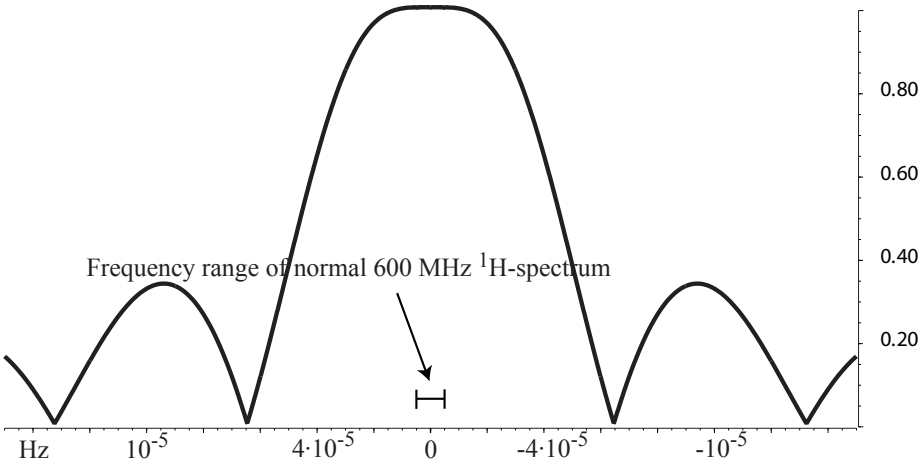


Figure 2.4: Simulated excitation profile of a 90° pulse. The figure shows xy-magnetization ($\sqrt{x^2+y^2}$) after a 90° pulse of length 15 μs. The values on the x-axis are the offsets from the frequency of the pulse, ν. Before the pulse M_z was 1. NMRSim was used to create the figure [38].

Equations 2.9 and 2.10 refers to relaxation parallel and perpendicular to the applied field, respectively.

The relaxation parallel to the applied field, which is called spin-lattice or longitudinal relaxation, is characterized by the time constant T_1. The name spin-lattice comes from the fact that the relaxation is associated with exchange of energy between the spins and environment, or the lattice. There are several different factors contributing to the relaxation, the most important are:

- dipole-dipole interactions
- chemical shift anisotropy
- spin-rotation interactions
- scalar coupling
- quadrupolar interactions
- paramagnetic compounds

For spin-$\frac{1}{2}$ nuclei the dipole-dipole interactions are the usually the most dominant. This type of interactions will be discussed later (see section 2.1.3.7 and 2.1.3.8). The chemical shift anisotropy, CSA, is another dominant factor. The shielding of the nucleus varies with the orientation of the molecule in the static magnetic field. Molecular tumbling will therefore give rise to a fluctuating magnetic field which again cause relaxation. If there are traces of paramagnetic compounds, either in the solution or in the glass of the NMR tube, this may be a dominating factor of relaxation. The other factors are of less importance.

For protons this relaxation is in order of a second. Due to the fact that these time constants does not vary much for proton in different bonding situations, proton time constants give little information and are consequently usually not measured. In addition the spectra are often too complex to allow relaxation measurements.

Relaxation measurements of ^{13}C are of greater importance. Because of dipole-dipole interactions the relaxation gets shorter the more protons are directly bonded to an atom. Relaxation may also tell something about the mobility of the different parts of a molecule.

The relaxation perpendicular to the field, which are called spin-spin or transverse relaxation, are of less importance. In addition to the factors affecting spin-lattice relaxation, the *true* T_2 is also affected by loss of phase coherence. This loss of phase coherence will cause signal intensity to be weakened. There is also a loss of phase coherence due to inhomogeneities in the magnetic field. During the spin-spin relaxation the energy of the system is not altered. Because of the additional sources of relaxation compared with spin-lattice relaxation, the spin-spin relaxation is always the fastest:

$$T_1 \geq T_2 \tag{2.11}$$

2.1.3.3 The FID and the Fourier transform

Once a pulse has been applied and we have achieved phase coherence, the precession of the nuclear spins will induce an oscillating electric current in the receiver coils. This is the NMR signal, or the free induction decay (FID). It is a decay because of relaxation.

This FID contains all the information about the nuclear spins, however, the human eye and brain are not able to interpret this information. A mathematical technique called Fourier transformation therefore has to be used:

$$S(\Omega) = \int_0^\infty s(t)e^{-i\Omega t}dt \tag{2.12}$$

This converts a function of time into a function of frequency. In reality an algorithm called fast Fourier transform (FFT) [39] is used because the FID is not a continuous, but a discrete function.

Because of the limitations of the current analog-digital converters, the frequency of the NMR signal has to be down-converted by the receiver. The output is the frequencies relative to the frequency of the RF pulse. It consists of two signals, which may be interpreted as the real and imaginary component of a complex signal. In addition to retaining the information about the sign of the relative frequencies, this makes it possible to phase correct the spectrum, and hence gain sensitivity. It also makes it possible to set the carrier frequency in the middle of spectrum. After digitalization and Fourier transformation the spectrum is phase corrected to yield the pure absorption spectrum (see Figure 2.6).

2.1.3.4 Signal averaging

In a single FID the noise is relatively dominating, hence the noise will also be dominating in the spectrum. When multiple FIDs are added the real signal will be added while noise partially will be averaged out. The result is that the signal-to-noise ratio will be improved by a factor $\sqrt{n}$, where *n* is the number of FIDs added.

2.1.3.5 Apodization

In order to enhance either sensitivity or resolution of a spectrum the FID may be multiplied/weighted by a function. This is know as apodization.

The noise of a FID is constant while the real signals are decaying. Concequently there is less noise compared to the real signals at the beginning of the FID. Thus, by

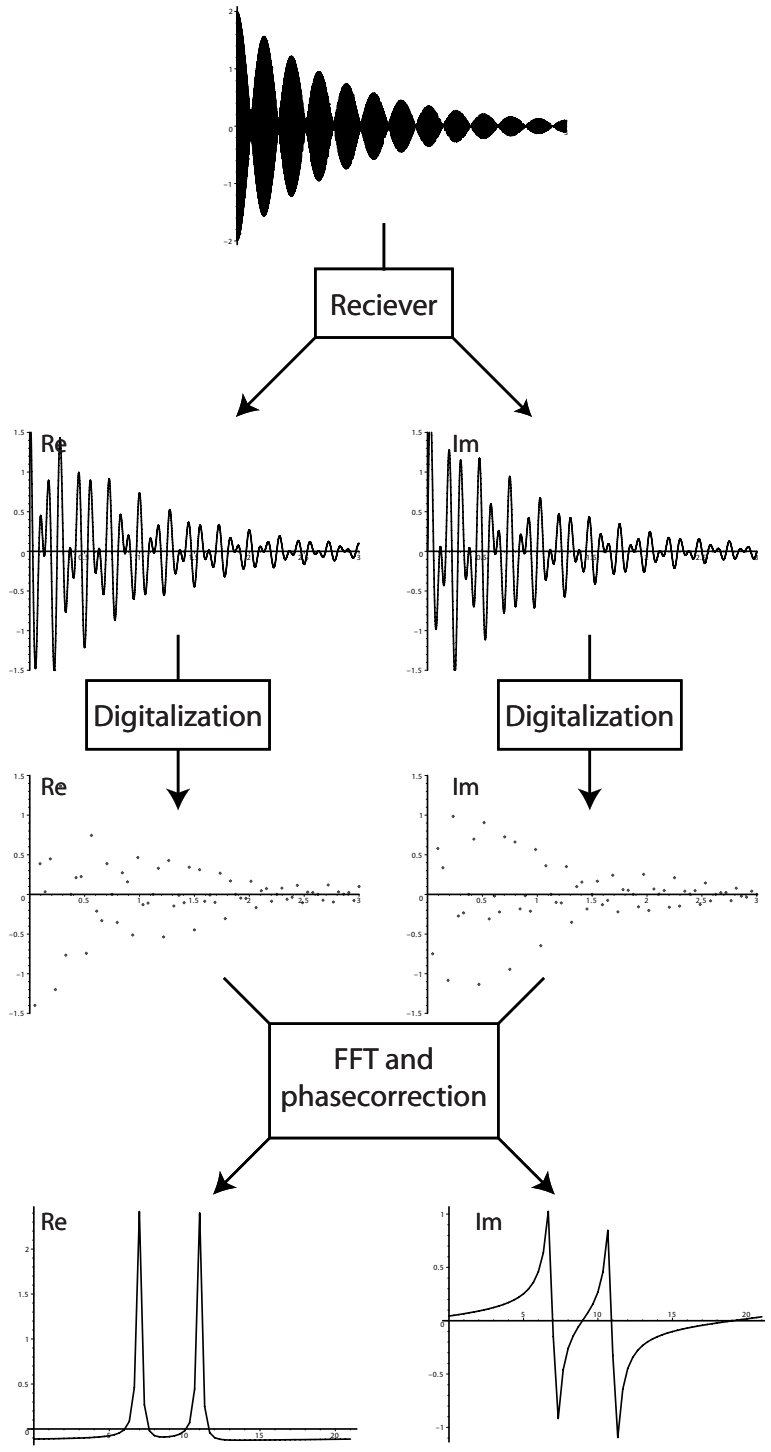


Figure 2.6: Fast Fourier transform of a FID consisting of two different frequencies with T_2=1 s. The receiver down-converts the frequencies and makes one real and one imaginary FID. These are then digitalized, Fourier transformed and phase corrected to yield a pure absorption spectrum.

applying an apodization function that enhances the first part of the FID, the signal-to noise-ratio will be enhanced. One example of such a function is the exponential decay function:

$$f(t) = e^{-\pi \cdot LB \cdot t} \tag{2.13}$$

where *LB* is the line broadening.

By enhancing the latter part of the FID it is possible to get better resolution at the cost of lower signal-to-noise ratio. This can be done with a squared sine function. It is most common to use a function that reaches a compromise between resolution and signal-to-noise ratio. The Lorentzian-to-Gaussian transformation is an example of this.

In multi dimensional NMR truncation is a major problem that will influence the choice of apodization function. An apodization function that suppress both the earliest and latest parts of the T1 FIDs is best to use. The Gaussian and squared sine functions are possible functions to use.

2.1.3.6 Indirect dipole-dipole coupling

A close examination of a 1D spectrum will reveal that some of the peaks have been split in to two line or more. This splitting is due to interaction with neighbouring nuclear spins. The interaction is mediated by electrons forming the chemical bonds, hence the name indirect coupling. Two nuclear spins that are coupled have equal splittings of the peaks. Because the splitting is not dependent on the magnetic field strength, it is measured in Hz rather than in ppm. The couplings are denoted ${}^{n}J_{ab}$ where *n* is the number of chemical bonds separating the nuclear spins *a* and *b* that are coupled.

Many 2D experiments are based upon these indirect couplings between nuclear spins. COSY and TOCSY are for instance based upon the homonuclear three bond coupling, while ^{1}H-^{15}N HSQC used in protein NMR is based upon ${}^{1}J_{\mathrm{NH}}$.

In addition to the obvious dependency on type of nuclei and the number of bonds separating them, the coupling constants are also dependent on several other factors. Most relevant in connection with structural determination is the dependency of ${}^{3}J_{\mathrm{HH}}$ on the dihedral angle, first discovered by Karplus [40]. This dependency is described by the Karplus equation:

$${}^{3}J = A\cos^2\phi + B\cos\phi + C \tag{2.14}$$

The coupling between the amid proton and the H^{α} proton in the backbone is especially important in structural determinations of proteins. Many groups have determined these coefficients, but the values used in this thesis are those determined by Pardi et al. [41]:

$${}^{3}J_{\mathrm{H}^{\alpha}\mathrm{H}^{\mathrm{N}}} = 6.4\cos^2\phi - 1.4\cos\phi + 1.9 \tag{2.15}$$

Other commonly used terms for the indirect dipole-dipole coupling are spin-spin coupling, scalar coupling and J-coupling.

2.1.3.7 Direct dipole-dipole coupling

In addition to the indirect coupling, there is also a direct dipole-dipole coupling. This coupling arises because nuclear spins affect each other, just like to bar magnets do. The size of the coupling depends on the nuclei involved, the distance separating them, and the angle between the vector joining them and the magnetic field:

$$d_{\mathrm{IS}} = b_{\mathrm{IS}}(3\cos^2\theta - 1) \tag{2.16}$$

where b_{IS} is the dipole-dipole coupling constant:

$$b_{\mathrm{IS}} = -\frac{\mu_0}{4\pi}\frac{\gamma_I \gamma_S \hbar}{r_{\mathrm{IS}}^3} \tag{2.17}$$

This coupling constant is in the order of kilohertz for two coupled protons. In isotropic liquids this coupling is averaged to zero, however, it forms the basis for the most important NMR parameter regarding structural determination - the NOE.

2.1.3.8 Nuclear Overhauser effect

The nuclear Overhauser effect was originally a technique to increase the intensities of some spins by irradiating others. Today the fact that NOE may lead to a decrease of signal intensity is also utilized.

The NOE may easily be visualized by looking at a system of two protons, I and S, that are not J-coupled. At equilibrium all transitions have equal population differences, and thereby equal signal intensities. We may now saturate spin S by irradiating it. At the beginning there will be no change in population differences for the I spin. However, relaxation will occur, mainly through two processes, W_0 and W_2, and this will change the intensity of I. If W_2 is the only process, at steady-state there will be a 50% increase in intensity for I:

$$f_{\mathrm{I}}\{\mathrm{S}\} = 0.5 \tag{2.18}$$

where $f_{\mathrm{I}}\{\mathrm{S}\}$ is the fractional increase of intensity for I when irradiating S. On the other hand, if W_0 is the only process, there will be a total loss of signal intensity:

$$f_{\mathrm{I}}\{\mathrm{S}\} = -1 \tag{2.19}$$

In general, the fractional increase of intensity for I when irradiating S is given by the equation:

$$f_{\mathrm{I}}\{\mathrm{S}\} = \frac{\gamma_{\mathrm{S}}}{\gamma_{\mathrm{I}}}\frac{W_2 - W_0}{2W_{1I} + W_0 + W_2} \tag{2.20}$$

In order to relax through the W_0 and W_2 process, fluctuations in the magnetic field at the appropriate frequencies are required. For the W_2 process, magnetic field fluctuations in the order of $2v_0$ are required, while low frequency fluctuations are required for W_0. The efficiency of the relaxation processes is dependent on the intensity of the fluctuations at the necessary frequencies, the spectral density, $J(\omega)$.

The necessary field fluctuations arise because of the combination of molecular tumbling and dipolar couplings with its angular dependency. Molecules tumbling quickly will thus generate fluctuations with high frequencies, making W_2 the most efficient relaxation process, while molecules tumbling slowly will generate low frequency fluctuation, making W_0 the most efficient relaxation process. According to Stokes' law the tumbling rate, or correlation time, τ_c, of a molecule is proportional to volume:

$$\tau_c = \frac{4\pi\eta_W r_H^3}{3k_B T} \tag{2.21}$$

where η_{W} is the viscosity of the solvent, r_H is the hydrodynamic radius and k_B is the Boltzmann constant. The consequence of this is that large molecules will have a negative NOE and small molecules will have a positive NOE.

Another possibility is to use a selective 180° pulse on spin S, instead of saturating it. The selective pulse is followed by a mixing time, τ_m, during which the NOEs will

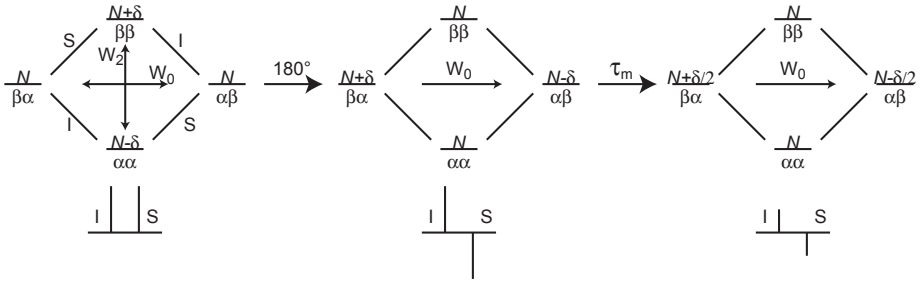


Figure 2.8: The changes in populations through the transient NOE experiment. S is the spin to which the 180° pulse is applied. No other factors than W_0 contribute to relaxation.

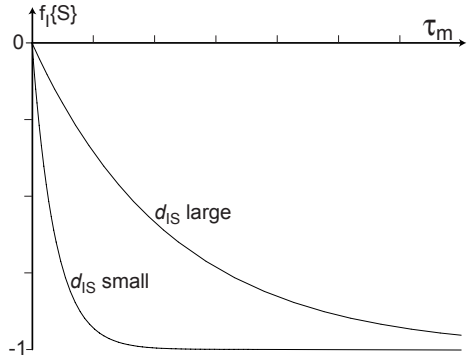


Figure 2.9: NOE buildup curve when relaxation only occur through W_0. The figure shows the enhancement, $f_I\{\mathrm{S}\}$, as a function of mixing time, τ_m.

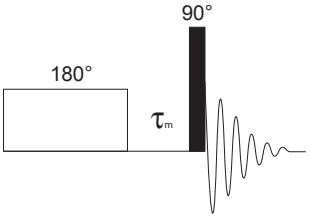


Figure 2.7: Pulse sequence for observing transient NOEs.

develop. The NOEs are turned into an observable signal by a 90° read pulse. This is the transient NOE experiment (see Figure 2.7). In contrast to the steady-state experiment the spin S is not continuously irradiated, and the populations will gradually relax back to their original populations. Figure 2.8 shows the populations trough a transient NOE experiment, and Figure 2.9 shows the build-up as a function of mixing time, τ_m.

The build-up of NOE is proportional to the square of the dipole-dipole coupling constant. Consequently, the NOE is has a strong dependency of distance. This build-up is characterized bye the cross relaxation rate constant, R_{cross}. When $\omega_0\tau_c \gg 1$ this is given by:

$$R_{\mathrm{cross}} = -b_{IS}^2 \frac{2}{5}\tau_c = \frac{C}{r_{IS}^6} \tag{2.22}$$

This relationship forms the basis for structural determination using NMR.

2.2 Solvent suppression

When studying a molecule by liquid state NMR, one serious problem often arises. The signal from the solvent totally dominates the spectrum, and the signals of interest are of poor quality, if even detectable. One simple solution to this problem is to use a deuterated solvent. However, because of the exchangeable amid proton this is not an acceptable solution when working with proteins. Three different solvent suppression methods will be used in this thesis.

2.2.1 WATERGATE

The WATERGATE (WATER suppression by GrAdient-Tailored Excitation) uses two magnetic field gradients separated by a 180° selective RF pulse, a gradient echo, to achieve water suppression. Resonances experiencing the 180° pulse will be rephased, while the other resonances will be dephased by the sequence. The original WATERGATE sequence [42] uses a hard 180°(x) pulse flanked by two soft 90°(-x) pulses to achieve the selective 180° pulse. In this thesis two different pulse trains have been used as the selective pulse. The simplest one, denoted 3-9-19 [43], consist of six pulses with length {3α, 9α, 19α, 19α, 9α, 3α} and phase {0°, 0°, 0°, 180°, 180°, 180°}. The other, W5 [44], uses the same scheme, but with 10 pulses instead of 6, optimized to produce the most uniform excitation profile. In addition, the sequence is run twice, with different gradient strength. The placement of the nil points and the width of the solvent suppression region are set by adjusting the delay between the pulses in the pulse train. A two-step phase cycle is used.

2.2.2 Gradients

In the gs-DQF-COSY experiment gradients are used to select the proper coherences, only the double quantum coherences are retained. Because water only have a single quantum coherence, the double quantum filter (DQF) implicitly provides water suppression. If the proton content in the water is too high, the phenomenon know as double quantum echo (DQE) makes it necessary to use magic angle gradients (see section 2.3.2).

2.2.3 Presaturation

The last method used for solvent suppression in this thesis is presaturation. This solvent suppression method is not as efficient as other methods, but it had to be used with COSY because it was impossible to use DQF-COSY with ACME (see section 2.5). Presaturation involves a weak RF field, which is employed during the recycle delay. This field will saturate the water signal, and ideally no water magnetization is present during the pulse sequence.

2.3 NMR experiments

2.3.1 WATERGATE

The 1D WATERGATE sequence consists of a simple 90° pulse followed by the sequence described in section 2.2.1 that dephases the water resonance (see Figure 2.10).

2.3.2 COSY

COSY was the first 2D experiment, originally proposed by Jeener in 1971, but developed by Ernst and co-workers [45]. It is the simplest, and maybe the most frequently used two-dimensional experiment. In its simplest form it consists of two 90° pulses, separated by an incremental delay, t_1 (see Figure 2.11 A). This incremental delay is common for all 2D experiments. In general, any two-dimensional experiment can be written in the following form:

$$preparation - evolution - mixing - detection$$

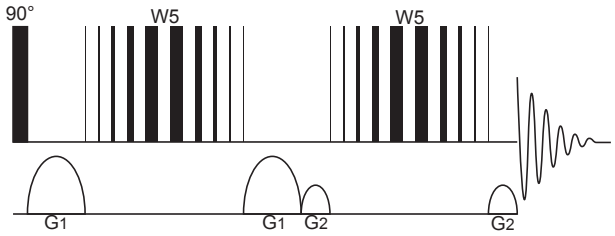


Figure 2.10: The 1D WATERGATE sequence with double gradient echo.

For the COSY experiment, the preparation consists of the first 90° pulse, which generates the single quantum coherences. During the evolution, the incremental delay, the resonances will rotate in the xy-plane with their corresponding frequencies. The orientation of the resonances at the end of the evolution period is a function of their frequencies and t_1, they will be frequency labelled. This frequency labelling is also responsible for the diagonal, which is common to all homonuclear 2D experiments. The last 90° pulse, the mixing, rotates the y-components (the x-components if it is a y-pulse) of the magnetization into the z-axis. This leads to a polarization transfer. Applying Fourier transformation to the FIDs leads to a normal spectrum, but with fluctuating peak intensities. The intensity of a peak is modulated according to its own frequency and the frequencies of spins to which it is scalar coupled. Applying a second Fourier transformation, with respect to t_1, gives a two dimensional spectrum, with cross peaks between spins that are scalar coupled.

Two different versions of the COSY sequence has been used in this thesis. The simples one is just the basic two-pulse COSY sequence with an eight step phase cycling to suppress axial peaks and quadrature glitches [46]. Water suppression is achieved by presaturation (see Figure 2.11 A).

The other experiment used was a double quantum filtered (DQF) COSY. Two different variants of DQF exist. The oldest one uses a third pulse and phase cycling to select the proper coherence pathway. The other variant of DQF, which was used here, employs gradients to select the proper coherence pathway (see Figure 2.11 B). With this method no phase cycling is necessary. Two magnetic field gradients and a 90° pulse constitute the DQ filter. The ratio between the two gradients is 1:2. These gradients will also generate a double-quantum echo (DQE). The magnitude of the echo has an angular dependency:

$$\frac{S_n}{S_\mathrm{I}} = \frac{0.25 t_2 (3\cos^2\theta - 1)}{2\mu_0 M_0} \tag{2.23}$$

where θ is the angle between the gradient vector and the magnetic field. By using gradients which have a magic angle with the z-axis, the DQE will be removed [47]. In connection with each gradient there is a spin-echo.

2.3.3 TOCSY

A COSY spectrum will only give cross peaks between spins that are linked by two or three bonds, with a few exceptions. Given a linear three-spin system, A-B-C, where A and C are not directly coupled, it is possible to modify COSY experiment (Figure 2.12 A) to also give an A-C cross peak. The modification consists of a spin-echo and a 90° pulse applied after the COSY sequence. This variant of COSY is called relayed-COSY (see Figure 2.12 B).

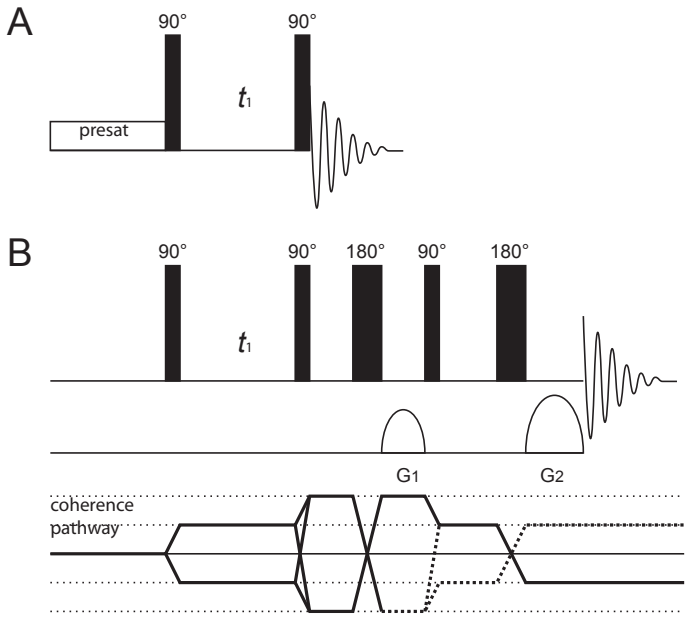


Figure 2.11: The two different COSY sequences used in this thesis. (A) The basic COSY sequence with presat used for solvent suppression. (B) gs-DQF-COSY.

It is possible to modify this sequence even more, to give cross peaks between all members of a spin system. This may be done by replacing the 180° pulse with a continuos RF field or a series of weak 180° pulses. This is the Total Correlation SpectroscopY (TOCSY) experiment [48]. The sequence of pulses is called an isotropic mixing sequence. During this sequence the magnetization is spin-locked along the y-axis. All resonances will have the same chemical shift. Because the coupling constants are larger than the chemical shift differences, coupled spins will be strongly coupled. Magnetization can now move among all the spins in a scalar coupled network. There are many isotropic mixing sequences. The one used here is called DIPSI-2 [49]. It consists of composite 180° pulses:

$$R = 320 \quad \overline{410} \quad 290 \quad \overline{285} \quad 30 \quad \overline{245} \quad 375 \quad \overline{265} \quad 370$$

arranged into $R\overline{R}\overline{R}R$ blocks. Values are in degrees and overbars indicates 180° phase shifted pulses. At the end of the sequence the water signal is suppressed by W5 with double gradients [44]. The pulse sequence can be seen in Figure 2.12 C.

When we are only interested in a small region of the TOCSY spectrum, a lot of time is spent on recording unnecessary information. If we could acquire only the F1 region of interest, the experiment would take significantly less time. This can be done by using BASHD-TOCSY [50]. BASHD-TOCSY uses double pulsed field gradient spin echo and selective pulses to frequency label only the desired resonances during t_1. The length and the frequency of the selective pulses determine which region is frequency labeled. The pulse sequence is displayed in Figure 2.12 D.

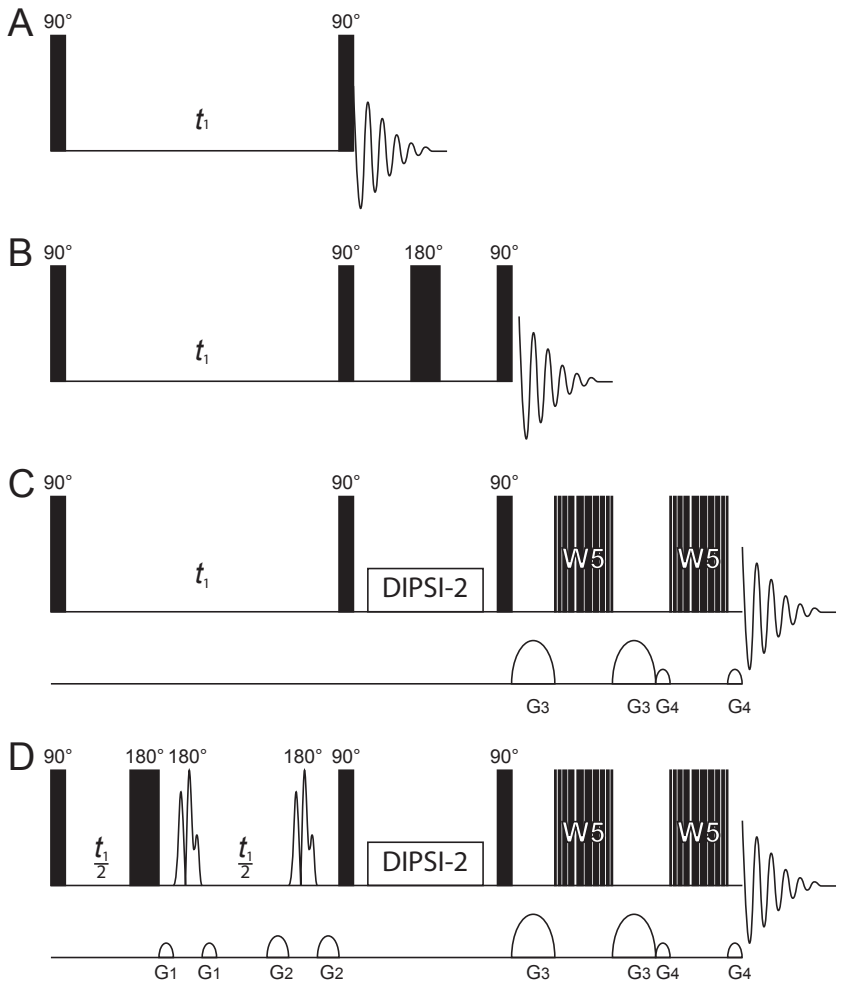


Figure 2.12: The figure shows the similarities between (A) COSY, (B) relayed-COSY, (C) TOCSY with double W5, and (D) BASHD-TOCSY.

2.3.4 NOESY

The pulse sequence for measuring transient NOEs (see Figure 2.13 A), can easily be extended to a two dimensional experiment. This is done by replacing the selective 180° pulse by two non-selective 90° pulses separated by an incremental delay, t_1, (see Figure 2.13 B) [51]. The result of this experiment is cross peaks between nuclei that are close in space. For an isolated two spin system, at short mixing times, the intensity/volume of the NOESY cross peaks are given by [52]:

$$a_{AB} \approx \frac{1}{2} M_0 (\frac{\mu_0}{4\pi})^2 \frac{\hbar^2 \gamma^4 \tau_c}{10 r_{AB}^6} \tau_m \propto \frac{\tau_m}{r^6} \tag{2.24}$$

This relationship applies when the molecule is large, and thereby tumbling slowly, $4\omega^2\tau_c^2 \gg 1$. In the presence of more spins, this simple relationship will become more complicated because of indirect NOE effects. The cross peaks may also be caused by chemical exchange.

During the incremental delay, t_1, and the NOESY mixing time, the water resonance will be affected by radiation damping. When using non-deuterated water, the large transverse magnetization of water induces a current in the transmitter/receiver coil. This current again generates a weak selective RF pulse that rotates the water resonance back to the +z axis faster than expected from natural relaxation processes. In order to have a reproducible water magnetization at the end of the mixing time, a gradient is used to dephase residual transverse water magnetization. This gradient also dephases transverse magnetization of the protein resonances, making phase cycling redundant.

The NOESY pulse sequence used in this thesis has an additional selective 90° pulse on the water resonance. This is applied immediately after the gradient. The phase is opposite of the following strong 90° pulse. As a consequence, the transverse water magnetization after the NOESY pulse sequence will be significantly reduced. This 90° pulse is known as a flip back pulse [53]. Residual transverse water magnetization is removed by the double WATERGATE sequence with W5 (section 2.2.1).

During t_1 the effects of radiation damping is minimized by the use of a bipolar gradient pair [54]. The entire NOESY pulse sequence used can be found in Figure 2.13 C.

2.3.5 HSQC and Fast HSQC

Multidimensional NMR experiments correlating ^{1}H and ^{15}N (HMBC and HSQC) has proven to be very useful in assigning chemical shifts in proteins. Of these two experiments, HSQC is the one generally preferred, among other factors, because of better resolution in the ^{15}N dimension.

The original HSQC sequence [55] consists of two INEPT sequences separated by the incremental delay, t_1 (see Figure 2.14 A). The first INEPT sequence will transfer magnetization from the high γ protons to the low γ ^{15}N. Frequency labelled ^{15}N z-magnetization will be generated by the incremental delay and the following 90° ^{15}N pulse. This magnetization is transferred back to the protons by the last INEPT sequence. By recording the spectrum in this manner sensitivity is increased, because the sensitivity is dependent on both the nucleus being excited and the nucleus being detected [56]:

$$\frac{S}{N} \propto \gamma_I \gamma_S^{3/2} \tag{2.25}$$

where γ_I is the excited nucleus and γ_S is the detected nucleus.

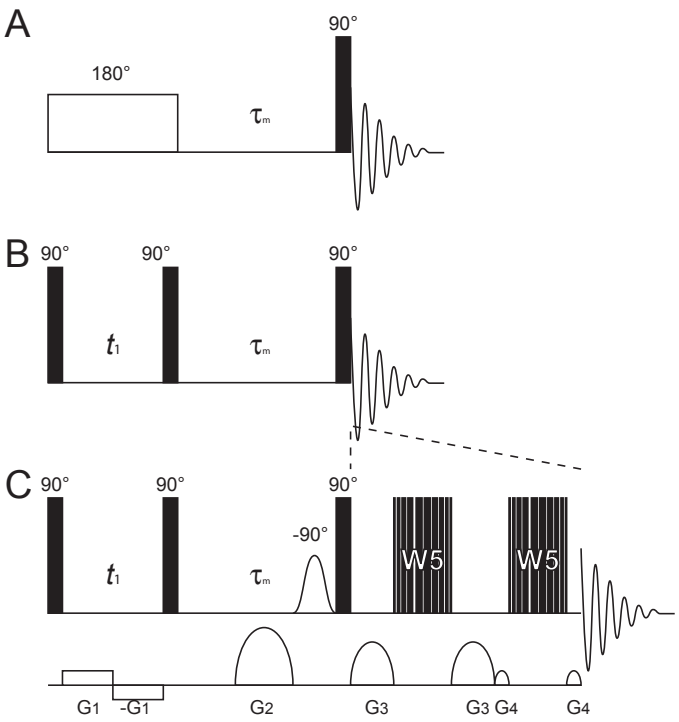


Figure 2.13: (A) Sequence for measuring transient NOEs. (B) The basic NOESY sequence. (C) The NOESY sequence used here.

The HSQC sequence used in this thesis has a few modifications. The last 180° proton pulse is replaced by the WATERGATE 3-9-19 solvent suppression [42, 43]. Two gradients, g_2 and g_3, are used to dephase unwanted transverse magnetization. The sequence also has a ^{1}H trim pulse at the end of the first INEPT sequence. The pulse program is shown in Figure 2.14 B.

In the Fast HSQC (FHSQC) [57] the magnetic field gradients during the incremental delay are equal in strength. This will ensure that no water magnetization remains at the start of the next scan. Any remaining magnetization would lead to loss of signal intensity for exchangeable protons. FHSQC also has two gradients of equal strength in the first INEPT sequence to suppress radiation damping [54]. Compared with WATERGATE-HSQC, FHSQC will give better signal intensities when using a short relaxation delay.

2.4 Assignment strategy

The different types of protons found in a protein are located in clearly defined regions in the spectrum. In the region from 6 ppm and higher the backbone amide protons are found. These protons have a fairly large chemical shift dispersion. The aromatic protons and side chain amine protons are also found in this region. The α-protons are found close to the water resonance, in the region from 3.5 ppm to 6 ppm. Below 3.5 ppm the aliphatic side chains are found, with the methyl groups being most upfield (see Figure 2.15).

The best method to use when assigning the ^{1}H resonances of an unlabeled protein is the *sequential assignment strategy* developed by Wütrich et al. (see for instance Protein NMR Spectroscopy [58]). The first step is to identify sets of H^N, H^α and aliphatic side chain resonances that belong to the same amino acid spin system using the scalar coupling information in TOCSY and COSY spectra. The aromatic protons are not scalar coupled to the rest of the amino acid. Information from a NOESY spectrum has to be used to assign these. Using the spin topology and a database of expected chemical shift for the different amino acids, an amino acid type are assigned to each spin system.

The best starting point for unravelling the spin systems is the fingerprint region, the H^α-H^N region above the diagonal (see Figure 2.15). Because of less distortion from the water resonance and higher F2 resolution this region is better than the region below the diagonal. Correlations with $H^N$are also a good way to determine the rest of the spin system. Correlations with H^α and side chain protons are used as a control and to assign ambiguous resonances.

The next step in the sequential assignment is to connect the amino acids. Because the $^4J_{\alpha\mathrm{N}}$ coupling constants connecting the residues are negligible NOE correlations have to be used. Intense NOEs from H^N, H^α or H^β on one residue to H^N on another usually correspond to two adjacent amino acids, respectively *i* and *i*+1. When enough residues are connected it is possible to match them to a unique part of the protein sequence.

If the aim of the assignment is to obtain a three-dimensional structure, the next step is to search for NOE correlations. The medium range NOEs characteristic for the α-helix ($d_{\alpha\mathrm{N}}(i,i+2)$, $d_{\alpha\mathrm{N}}(i,i+3)$ and $d_{\alpha\mathrm{N}}(i,i+4)$) and the β-sheets are relatively easy to obtain using information from CSI and $^3J_{\alpha\mathrm{N}}$. Long-range NOEs are considerably more difficult to assign, but automation (see section 2.7) and x-ray structures of homologous proteins may aid this process.

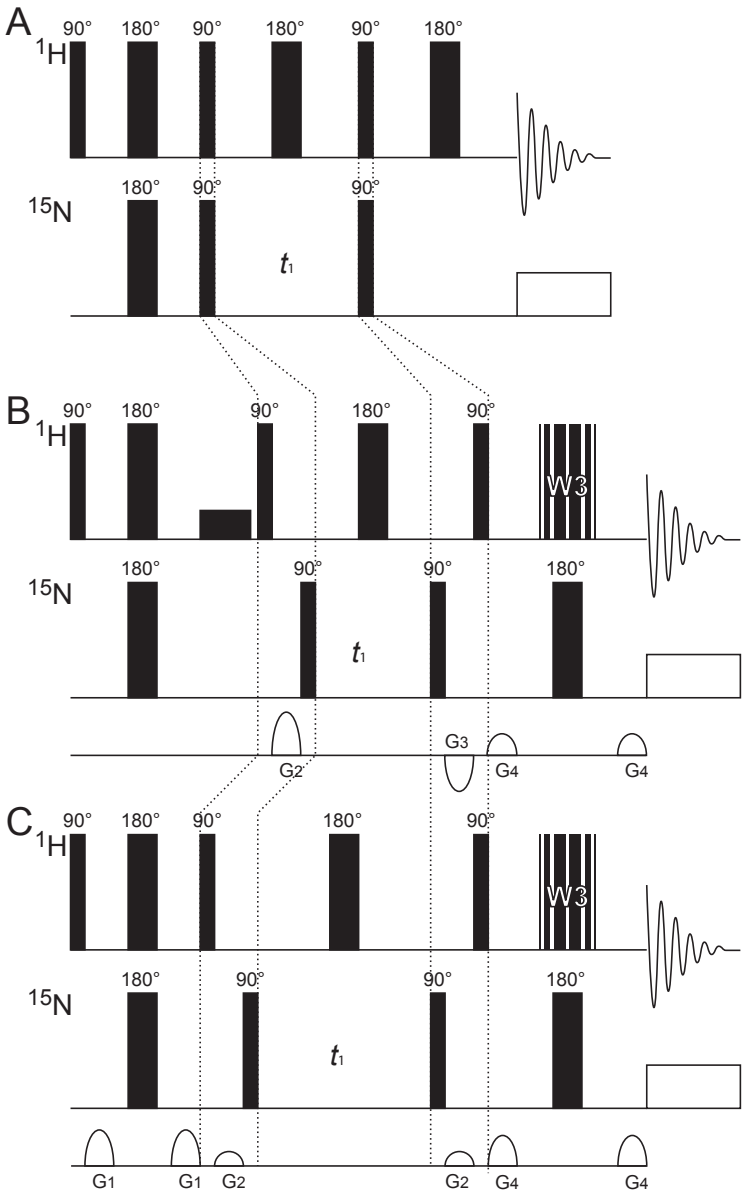


Figure 2.14: (A) Original HSQC, (B) the modified HSQC sequence used here, and (C) FHSQC. The figure tries to highlight the differences and similarities between the three sequences.

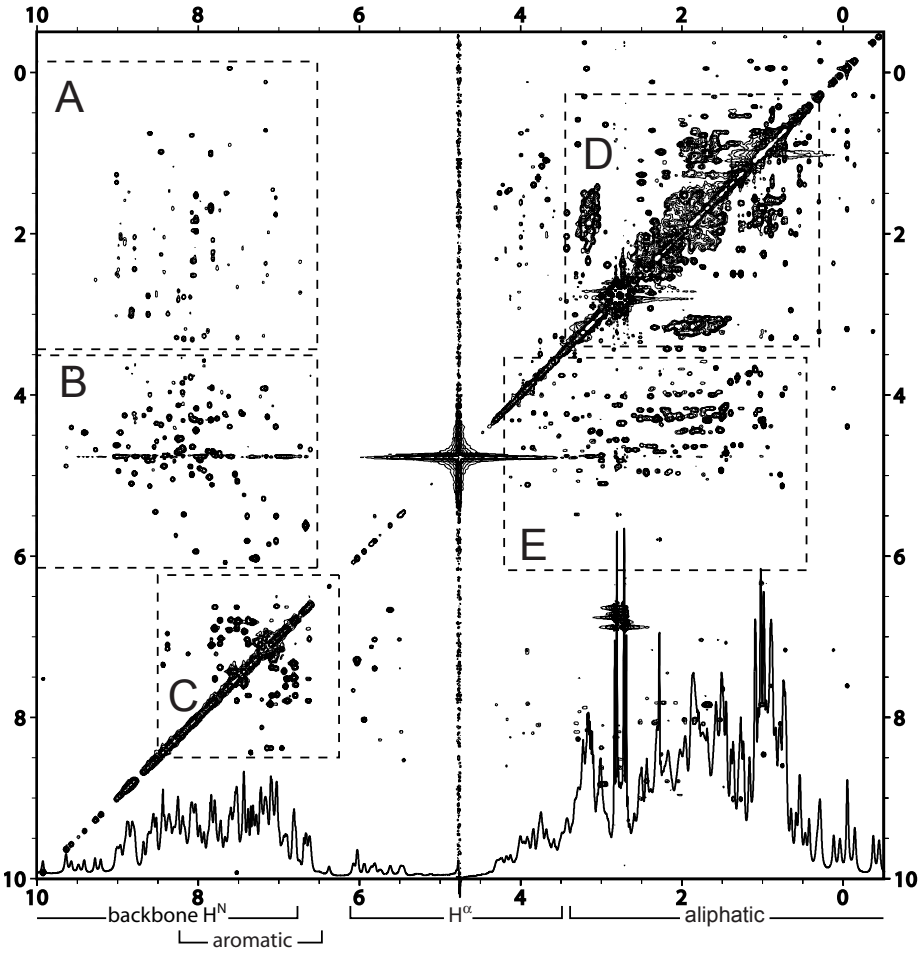


Figure 2.15: The different regions of a TOCSY spectrum of a protein (BLA). A 1D spectrum is shown. (A) H^N-aliphatic correlations, (B) fingerprint region, (C) aromatic cross peaks, (D) aliphatic cross peaks, and (E) H^{α}-H^{β}. To simplify the spectrum, only the strongest cross peaks are shown.

2.5 ACME

The ^{1}H-^{1}H J-coupling constants provide information that is very important regarding structural determination (see section 2.1.3.6). Several different methods have been developed for measuring the ^{1}H-^{1}H J-coupling constants. These methods include direct measurement from COSY spectra and various homo- and heteronuclear methods based on E.COSY. They are, however, not satisfactory when working with proteins at natural abundance where signals overlap.

There is one major problem when measuring coupling constants from regular COSY spectra. The coupling constant is a function of the linewidth and the intensity of the signal, in addition to the passive and active couplings. By constraining the multiplet intensity, it is possible to extract coupling constants from single multiplets or small groups of overlapping multiplets without knowing details about the spin system . This method is known as amplitude-constrained multiplet evaluation (ACME), and is performed by the software package ACME [59], running as a part of NMRPipe [60].

The ACME method relies on the fact that the intrinsic intensity for all multiplets is the same, and identical to that of the diagonal signals. The intensity of the diagonal peaks is determined by an diagonal-only spectrum. Separable signals are fitted to give the intrinsic intensity.

2.6 CSI

It has long been known that there is a clear relationship between the chemical shifts and the secondary structure in proteins. More specifically, in an α-helix the H$^\alpha$ proton will have an upfield chemical shift with respect to the random coil value, while in a β-sheet the shift will be downfield [61].

Wishart et al. have developed a simple, yet efficient, method for assigning the secondary structures in proteins using only the H$^\alpha$ shift [62]. Each residue is given an index, "-1" , "0" or "+1", depending on its H$^\alpha$ chemical shift relative to the random coil value. If the chemical shift is less than the range indicated in Table 2.1 the residue is given the index "-1". If it is larger, it is given the index "+1". Otherwise it is given the index "0".

Table 2.1: Chemical shift ranges used with CSI

residue	H$^\alpha$-range	residue	H$^\alpha$-range
Ala	4.35±0.10	Met	4.52±0.10
Cys	4.65±0.10	Asn	4.75±0.10
Asp	4.76±0.10	Pro	4.44±0.10
Glu	4.29±0.10	Gln	4.37±0.10
Phe	4.66±0.10	Arg	4.38±0.10
Gly	3.97±0.10	Ser	4.50±0.10
His	4.63±0.10	Thr	4.35±0.10
Ile	3.95±0.10	Val	3.95±0.10
Lys	4.36±0.10	Trp	4.70±0.10
Leu	4.17±0.10	Tyr	4.60±0.10

Local densities of "-1" and "+1" indicate α-helix and β-sheet respectively. A group of four or more "-1's", not interrupted by a "+1", designates an α-helix. Three or more

"+1's", not interrupted by "-1", designates a β-sheet. A group may include "0", but not more than 30%. Other regions are designated as random coils.

Secondary structures determined by CSI are not 100% correct. An accuracy of 90-95% is typical.

2.7 Structure calculation using ARIA

Calculating a protein structure from experimental data is all about finding the lowest energy structure that is consistent with the experimental data [63]. Several stages are involved in this search. The first step is the creation of starting structures using the covalent structure and the experimental restraints. The next step is Restrained Molecular Dynamics (RMD). This involves including the experimental restraints in a molecular dynamics simulation. The result is low energy structures. In order to overcome local minima and find the global minimum, a temperature increase of the system i simulated, followed by slow annealing. This last step is called simulated annealing.

ARIA (Ambiguous Restraints for Iterative Assignment) is a structure calculation software that integrates automated NOESY assignment with structure calculation [64]. The NOESY assignment is done in an iterative way using ambiguous distance restraints (ADRs) [65, 66]. In principle the user only needs to provide a list of chemical shifts, an unassigned NOESY cross-peak list, and obviously the protein sequence. The chemical shifts list and NOESY peak list may be files directly taken from various spectra assignment programs (among others Sparky, ANSIG, NMRView and XEASY). Other types of experimental data may also be included in the calculation. The software package CNS (Crystallography & NMR System) [67] takes care of the molecular dynamics and simulated annealing.

2.7.1 NOE restraints

As a consequence of the large number of protons in a protein, several protons will have the same chemical shift. This makes it very difficult, or even impossible, to assign the NOESY cross-peaks, as they may be the result of several different interactions. Ambiguous NOESY cross-peaks were earlier not included in structure calculations. By using iterative assignment strategies some of these cross-peaks may be assigned based upon preliminary structures. The introduction of ADRs has made it easy to incorporate these restraints directly into the structural calculations in an automated fashion.

The first step in the structural calculations is the calibration of the NOESY cross peaks. In ARIA this is done using the *S* lowest energy structures from the last iteration step. *S* is typically 7 out of the 20 calculated structures. The mean distance between each proton pair is calculated from these structures:

$$\hat{d}_{ij} = \frac{1}{S}\sum_{s=1}^{S} d_{ij,s} \tag{2.26}$$

For iteration 0, the distances from the template structure is used. All the proton pairs, with d_{ij} less than 6 Å and which has an fully assigned NOESY cross-peak, are then used to calculate the calibration factor, *C*. The distances observed from the NOESY cross-peaks, d^{obs}, are then calculated as follows:

$$d^{obs} = (CV)^{-\frac{1}{6}} \tag{2.27}$$

where *V* is the volume or intensity of the NOESY cross-peak.

For each unassigned NOESY cross-peak an ADR is now created. An effective distance, $\overline{D}$, containing contributions from all pairs of protons that are possible assignments of the NOE, is defined:

$$\overline{D} = \left(\sum_{a=1}^{N_\delta} d_a^{-6} \right)^{-\frac{1}{6}} \tag{2.28}$$

where the index *a* runs over all N_δ possible assignments of the NOESY cross-peak. The ADR is now created by demanding that $\overline{D}$ stays within a specified range from d^{obs} calculated from the cross-peak. This restraint is now used in the molecular dynamics simulation.

Within the specified range from the calibrated distance, d^{obs}, the energy contribution from the NOESY cross-peak is defined as 0. Outside this range, but close to the calibrated distance, it is given a harmonic-like energy potential. When $\overline{D}$ is much larger than the calibrated distance the potential is linear. For a detailed description, see the article by Nilges and Donoghue [68].

For each iteration a fraction of the most unlikely contributions to an ADR is removed. In this way the NOESY cross-peaks may in the end be fully assigned.

2.7.2 Other experimental data

With ARIA it is also possible to use other types of experimental data. These data include torsion angles, scalar couplings, residual dipolar couplings (RDCs), hydrogen bonds, disulphide bridges and planarity restraints. Output files directly from the CSI software [69] may also be used. These additional restraints are especially important in α-helical regions, where the ambiguity in the NOESY cross-peaks is large.

2.7.3 Final calculations

After the last iteration a water refinement is done one the 7 structures with lowest energy. The energies, root mean square deviations (RMSD) values and violations of experimental restraints are calculated.

2.8 pK_a determinations

It has been show that the protonation of the aspartate and glutamate residues in BLA plays an essential role in its binding to a membrane (see section 1.2.6). These residues have to be protonated before BLA can interact with a negatively charged membrane. On this basis it would be interesting to know the individual pK_a values of the aspartate and glutamate residues. The pK_a values can also tell us something about the calcium binding properties (see section 1.2.2).

When a side-chain carboxyl group is deprotonated the electron densities around the nearby atoms is altered. Compared with the deprotonated state, the protons close to the carbonyl group will have a lower electron density when it is protonated. This will lead to a change in the chemical shift. The side chain protons in aspartate and glutamate will have a chemical shift that is in the order of 0.2 ppm lower when carboxyl group is deprotonated [70].

These pH dependent chemical shifts, both for protons and ^{13}C, have been used by many groups to determine the pK_a of the side chain carbonyl groups, in order to study, among others, ligand binding [71, 72], catalysis [73] and protein stability [70, 74]. The chemical shift is fitted to a modified Hill equation [75] to determine pK_a:

$$\delta_{\mathrm{obs}} = \delta_{\mathrm{A}}[1 + 10^{n(\mathrm{p}K_a - \mathrm{pH})}]^{-1} + \delta_{\mathrm{HA}} \left\{1 - [1 + 10^{n(\mathrm{p}K_a - \mathrm{pH})}]^{-1}\right\} \qquad (2.29)$$

where δ_{obs} is the observed chemical shift at the given pH, and δ_{A} and δ_{HA} is the chemical shift when the carboxyl group is fully deprotonated and fully protonated. *n* is the Hill coefficient. Ideally this coefficient should be 1, but it may be altered by different factors. A Hill coefficient less than or greater than one suggests that the pK_a is changing while the carbonyl group is titrated. If the pK_a is greater at high pH than at low pH the Hill coefficient will be less than one. This effect is likely to be caused by other neighbouring carboxyl groups or histidine residues with low pK_a values [70, 76].

Chapter 3

Experimental

3.1 Sample preparation

3.1.1 Washing of NMR tubes

To achieve good NMR spectra, it is essential that the NMR tubes are clean. The NMR tubes were first rinsed thoroughly using distilled water, then washed using an ultrasonic bath (Ultraschallgerät T 460/H from Elma). The water in the bath contained a few milliliters of Deconex, the temperature was 50 °C. The tubes lay in the bath for at least 30 minutes. After exchange in EDTA for another 30 minutes, the tubes were rinsed thoroughly first using distilled and then using deionized water. Finally the tubes were dried using a vacuum pump. If a NMR tube did not appear clean afterwards, it was washed in concentrated nitric acid, then the entire washing procedure was repeated. New NMR tubes were also cleaned before use.

3.1.2 Protein sample

Bovine holo-α-lactalbumin was obtained from Sigma-Aldrich (prepared from raw, unpasteurized milk). Approximately 14 mg protein was dissolved in 550 μl phosphate-citric acid buffer and 50 μl D_2O. This corresponded to approximately 1.6 mM BLA. The phosphate-citric acid buffer consisted of 5 mM citric acid, 10 mM Na_2HPO_4 and 0.1 M NaCl. pH of the pure buffer was 4.5. After adding the protein, pH was adjusted to approximately 5.6 using a diluted NaOH solution. All solutions were made using super distilled water and analytical grade reagents. DSS were added to some of the samples for use as a chemical shift reference. Except for Na_2HPO_4, which was obtained from J. T. Baker, all chemicals were obtained from Sigma-Aldrich. 550 μl of this protein solution was transferred to a clean Wilmad 528-PP NMR tube.

3.2 NMR acquisition and processing for assignment

All the NMR spectra used in this thesis were acquired using a Bruker Avance DRX 600 NMR spectrometer and a TBI probe head with x-, y- and z-gradients. The temperature of the sample was 314.0 K. The probe was shimmed using the program *gradshim* and the shim group *shim27,* which includes all the shim functions of the probe. The basis for the shimming was usually the shim settings from the last "run".

All the high power proton pulses were calibrated with the pulse program *zg* and the 360°-method. The chemical shift of the water resonance was determined at the same time.

The pulse programs used can be fund in the appendix.

3.2.1 1D/WATERGATE

A WATERGATE spectrum was acquired before every 2D spectrum. This was done in order to check that sample was in proper shape and that the water suppression scheme was working.

3.2.1.1 Acquisition

The pulse program used was *dpfgsew5*. The spectra were acquired with 32k points in the FID (TD), 32 scans (NS), 16.9 ppm spectral width (SW) and the receiver gain (RG) set to 128. The delay for binomial water suppression, D19, was set to 90 μs. This both ensured that the water signal would be completely suppressed and that protein signals at the highest and lowest chemical shift would not loose any intensity. The experiment time was 2 minutes.

3.2.1.2 Processing

The FIDs were multiplied by an exponential function using a linear broadening (LB) of 5 Hz. They were then Fourier transformed and phase corrected to yield pure absorption spectra.

3.2.2 COSY

3.2.2.1 Acquisition

The pulse program *mqsgs1d* with the gradient program *2sinexyz* was used to calibrate the magic angle gradients and to calculate the receiver gain. The COSY spectrum was then acquired with the pulse program *cosygsmftp*. The gradient program *2sinexyz* together with the calibrated magic angle gradients suppressed the water signal. The spectrum was acquired with 640 and 4k points in respectively F1 and F2. 56 scans were recorded for each increment. The spectral width was 14.98 ppm in both dimensions.

3.2.2.2 Processing

A squared cosine apodization function (WDW=qsine, SSB=2) was used in both dimensions. Forward linear prediction to 1k points was used in the F1 dimension. The spectrum was Fourier transformed into 2k points in both dimensions. The spectrum was phase corrected before it was converted to the UCSF format and analysed in Sparky [77].

3.2.3 TOCSY

3.2.3.1 Acquisition

The pulse program *dipsi2stdpfgsew* was used in the TOCSY experiment. The delay for the water suppression, D19, was the same as in the 1D spectra, 90 μs. The spectrum

was recorded with 1k points in F2, 768 points in F1 and 64 scans. The TOCSY mixing time was 80 ms and the power level for the TOCSY-spinlock was 9.54 dB (compared with 0.00 dB for the high power pulses). The sweep width was set to 11.97 ppm. At a later point a supplementary spectrum was recorded with a sweep width of 14.985 ppm. This was done because the first spectrum did not cover all the signals.

3.2.3.2 Processing

A squared cosine window function (SSB=2, WDW=qsine) was used in both F1 and F2. The FIDs were Fourier transformed into 2k points in both dimensions. Forward linear prediction to 2k points was used in F1. No linear prediction was used in F2. The spectrum was phase corrected so that all peaks were positive, and was finally converted to the UCSF format for analysis in Sparky.

3.2.4 NOESY

A series of NOESY spectra were recorded in order to determine the structure of BLA.

3.2.4.1 Acquisition

All the NOESY spectra were recorded just after the first TOCSY spectrum, with the same sample, temperature, pulse length and shimming. Between each 2D spectrum a 1D spectrum was recorded to verify that the sample was all right. The pulse program used for the NOESY experiments was *noesyfprsdpw5*. This experiment too uses a double W5 for solvent suppression. The spectra were recorded with 576 and 4k points in respectively the F1 and F2 dimension. Both dimensions covered 14.985 ppm. The number of scans was set to 64. The 90° shaped pulse used was a Gaussian pulse, *gaussnf.1k*, with length 10 ms. The power level of the shaped pulse was calibrated using the XWIN-NMR shape tool. The spectra were recorded with the following NOESY mixing times in random order: 30 ms, 60 ms, 90 ms, 120 ms and 150 ms. Except for the mixing time the parameters were the same for all the spectra.

3.2.4.2 Processing

All the spectra were processed with exactly the same processing parameters. They were Fourier transformed into 2048 points in both dimensions. The window functions used was a squared cosine function (WDW=qsine, SSB=2). Forward linear prediction to 1k points was used in F1. After the spectra were phase corrected they were base line corrected in both dimensions and converted to the UCSF format.

3.2.5 ^{1}H-^{15}N HSQC versus FHSQC at natural abundance

Two spectra were acquired to compare the HSQC with the new and improved Fast HSQC (FHSQC) (see section 2.3.5). All acquisition and processing parameters were exactly the same. The spectra were recorded consecutively, without removing the sample from the spectrometer. A 1D spectrum was recorded before and after each HSQC spectrum.

3.2.5.1 Acquisition

The pulse program used for regular ^{1}H-^{15}N HSQC was *invif3gsst19*, for the Fast HSQC it was *fhsqc.nf*. Both spectra were acquired with 2k points in F2, 64 points in F1 and 800 scans. The spectral with was 11.02 ppm for the proton dimension and 28.00 ppm for the ^{15}N dimension. The delay for the water suppression was 126.6 μs, this corresponds to 3949 Hz or 6.58 ppm. The relaxation delay was set to 1.2 s. The Calibration of the 90° ^{15}N was done by comparing the calibrated 90° proton pulse to the 90° proton pulse of a sample with a known 90° ^{15}N pulse. A linear correlation between the proton and ^{15}N pulse was assumed. The acquisition time was equal for the two variants.

3.2.5.2 Processing

Both spectra were Fourier transformed into 2048 points in both directions. A squared cosine apodization function was used for both dimensions. Forward linear predictions to 256 points was used in F1. After processing the spectra were compared using Sparky.

3.3 Assignment work

The assignment of the fingerprint region done by Forge et al. [31] and Halskau et al. [2] were used as a starting point for the assignment work. Because of different experimental conditions it was not possible to copy the assignments directly. In addition the assignment by Forge et al. was done on the recombinant protein, which has a valine instead of a methionine in position 90. The assignments were also checked with the assignment done by Alexandrescu et al. [27].

Unambiguous H^{α}-H^{N} cross peaks were used as entry points. H^{α}-H^{N} cross peaks which had only one possible assignment were considered as unambiguous if they had a ^{15}N chemical shift that agreed with the literature. A combination of the TOCSY, COSY and NOESY spectra were used to derive the rest of the spin system and any possible aromatic protons belonging to the residue. The chemical shifts of the side chain were used as a quality assurance for the identity of the residue.

The NOESY spectra were used to "walk" down the polypeptide chain in both directions as far as possible. The literature and the chemical shifts of the side chain were used to confirm these residues.

The NOESY spectra were also searched for cross peaks that could be used to determine the secondary structure. This applied especially to the characteristic $d_{\alpha N}(i, i+3)$ found in α-helices and long range connectivities in the β-sheets. The assigned cross peaks, and all other clearly defined cross peaks in the 150 ms NOESY spectrum were integrated for use in the structural calculations. A software package, AUTOPSY (AUTOmated Peak picking for NMR SpectroscopY) , was also used to pick and integrate the NOESY cross peaks [78]. Because of different methods of integration the manually and automatically integrated spectra were treated as two different spectra.

3.4 ACME

The gs-DQF-COSY spectrum acquired for the assignment work was not suitable for use with ACME. A new, regular COSY spectrum had to be acquired, with presaturation used as solvent suppression.

3.4.1 Acquisition of COSY spectrum

The pulse program *zgpr* was used to calibrate the power level for the presaturation. The COSY spectrum was acquired using the pulse program *cosyprsh*. The time domain was set to 4k in the F2 dimension and 400 in F1. The number of scans was 32 and the sweep width 11.02 ppm. In order to regain full signal intensity for all resonances between each scan the relaxation delay was set to 5 seconds.

3.4.2 Processing

The data was first converted to a format that NMRPipe was able to read. A cosine apodization function was used for both dimensions. The F1 dimension was zero filled to 2k points. The spectrum was then phase corrected yielding a regular COSY spectrum with antiphase cross peaks. The diagonal was removed using a macro. A script took care of all this. Another script created a diagonal-only spectrum.

3.4.3 Estimation of intensity

The diagonal-only spectrum was used to estimate the intrinsic intensities of the cross peaks. Due to extensive overlapping most resonances could not be used to estimate the intensities. Only 18 resonances in the amid region, between 5.5 ppm and 10.1 ppm, were found suitable. The intensities were determined by starting ACME in Diagonal Fitting Mode. Individual diagonal peaks were selected and fitted to determine their intensities. Where the diagonal peaks were partly overlapping multiple peaks were selected and fitted. The variables that were fitted was width, intensity and location of the peaks. The other variables were fixed at zero (see Figure 3.1). The intensity used in the measurement for the coupling constant was an approximate average of the estimated intensities.

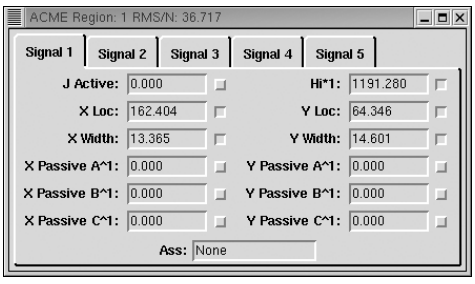


Figure 3.1: The parameter window in ACME. When determining the intrinsic intensity, the coupling constants, both the active and the passive, are kept at zero.

3.4.4 Measurement of coupling constants

The COSY spectrum, with the diagonal removed, was loaded into ACME. Where cross peaks were clearly separated from other, individual peaks were fitted. The intensity was kept fixed at 1000. Cross peaks in the H^{α}-H^{N} region have no passive X couplings, therefore this was kept fixed at zero. The initial estimate of both the active coupling and the passive Y coupling was 15 Hz. Where cross peaks were overlapping, multiple

peaks were fitted simultaneously. The effect of different intensities on the coupling constant was also examined for residues D87 and K94.

3.5 CSI

Wishart et al. have made a small computer program, CSI, that can determine the secondary structure based on H^{α}, C, C^{α} and C^{β} [69]. This program is currently available in version 2.0. Instead of using this program to determine the secondary structure a spreadsheet was used. The H^{α} chemical shift values from the assignment work were put into this spreadsheet. For the glycine residues the average of the two chemical shifts were used to calculate the chemical shift index. The rules given by Wishart et al. [62] were used to determine the secondary structures.

3.6 Structural calculations

3.6.1 Initial ARIA setup

The initial step of the structural calculations consists of preparing the input files, and specifying the location of these files. Running ARIA for the first time will create the directory tree, with the run.cns file, and convert the input files to a CNS readable format.

3.6.1.1 NOE and chemical shift list files

ARIA does not use NOE build-up curves from multiple mixing times to calculate the structure. Only a single mixing time needs to be used. The spectrum with 150 ms mixing time was used. As a compensation ARIA uses a spin diffusion correction. ARIA is able to read both peak lists and chemical shift lists exported from Sparky. The peak lists from Sparky were exported with the following columns: “Assignment”, “w1”, “w2”, “Volume” and “Fit Height”. Because of different methods of integration the two peak list files were not combined to one file. Some resonances also had to be renamed so that ARIA would interpret them correctly. For instance: the two β-protons in D46 were named from D46HB’ and D46HB” to D46HB*, because it was unknown which was $H^{\beta 2}$ and $H^{\beta 3}$. The peak lists had the following format:

```
      Assignment     w1      w2        Volume  Fit Height

              ?-?  1.167   7.228  8.98e+006 lo     362571
     A40HB*-G51HN  1.449   9.485  2.25e+007 lo     159043
     A40HB*-L52HA  1.449   4.045  4.08e+007 lo     303514
         E25HG*-?  0.115  10.480  1.22e+007 lo     238188
    W104HH2-L96HN  6.955   9.208  8.63e+006 lo     147917
```

and the resonance list file had this format:

```
     Group   Atom  Nuc     Shift   SDev  Assignments

        T4     HA   1H     4.806  0.003      3
        T4     HN   1H     7.976  0.003      4
        C6    HB*   1H     3.246  0.000      1
        C6    HB*   1H     3.205  0.002      2
```

The peak list from the assignment work consisted of 920 fully assigned peaks and 831 unassigned peaks. The peak list generated automatically with AUTOPSY consisted of 3713 unassigned peaks.

3.6.1.2 CSI

Residues that according to CSI were located in an α-helix were given a ϕ-angle of -57° and a ψ-angle of -47°. There were found no NOEs proving that the residues V99-I101 and L110-E113 were involved in a β-sheet. On this basis the dihedral restraints from CSI for these residues were removed. The restraints involving glycine were also removed. The NOE connectivities show that the three β-strands found from I41 to I59 are oriented antiparallel (see Figure 4.6 on page 64). These residues were thus given the ϕ-angle -139° and the ψ-angle 135°, the values for the antiparallel β-sheet. Outside of a range $\pm 10^\circ$ the dihedral angles derived from CSI were given a harmonic potential. The number of dihedral angle restraints from CSI was 130.

3.6.1.3 $^3J_{\mathrm{H^N H^\alpha}}$ coupling constants

ARIA can also derive dihedral angles from $^3J_{\mathrm{H^N H^\alpha}}$ coupling constants. The coupling constants determined with ACME were included in the structural calculations with an uncertainty of 1 Hz. Residues with coupling constant between 6.0 Hz and 8.0 Hz were not included, because these coupling constants may reflect motional averaging. The coefficients determined by Pardi et al. [41] were used by ARIA to derive the dihedral angles. The number of coupling constant restraints included in the calculations was 44.

3.6.1.4 Ca-restraints

Ca^{2+}-O restraints were introduced so that the calcium ion would be correctly placed in the structure. The distance to the three carboxyl and two carbonyl oxygens known to participate in the ligand binding (see section 1.2.2) were restrained to 2.4 ± 0.2 Å. This is in agreement with the Ca^{2+}-O distances in calculated for different EF-hand proteins [79].

3.6.2 Editing of run.cns

The run.cns file contains all the parameters necessary for the structural calculations. Only a few changes were made in this file, most of the parameters were kept to their default values. Based on the homology to c-type lysozymes, the correlation time of BLA was set to 10 ns [80]. The spectrometer frequency was also specified. Because the calcium ion was included in the calculation, Cartesian dynamics had to be used in the molecular dynamics simulations. In order to include the calcium ion in the automatically generated topology and template files had to be replaced.

3.6.2.1 Preparation of topology and template files

The topology and template files were based upon an x-ray structure of BLA, 1F6S.pdb [4]. Because 1F6S.pdb is a x-ray structure all protons and some side chain atoms were missing. These were added using the WHAT IF Web Interface [9]. Protons were also added to the three histidine residues to make them positively charged, since the experiments were carried out at pH 5.6. The nomenclature of the proton was then changed to fit the IUPAC nomenclature [3]. The exception was glycine. For some reason ARIA does not use the IUPAC nomenclature for glycine. The names used for the protons in glycine are HA1 and HA2.

The edited 1F6S.pdb-file was then used to generate the “begin”-files: 1F6S.psf, 1F6S_template.pdb, methyls.tbl and setup_swap_list.tbl. The calcium ion was then

manually added to the template- and topology-file. The automatically generated files were replaced by these new files.

3.7 pK_a determinations

The pK_a values for all the aspartic and glutamic acid residues were determined by plotting the proton chemical shifts as a function of pH. The protons closest to the carboxyl group were used. In the cases where this chemical shift was unknown, other protons were used. The chemical shifts of the relevant residues are found in Table 3.1.

Table 3.1: Chemical shifts for the aspartic and glutamic acid residues

	Resonances					
Residue	H^N	H^α	H^{β}'	H^{β}"	H^{γ}'	H^{γ}"
D14	9.17	4.57	2.46	2.55		
D37	7.08	5.17	2.49	3.16		
D46	8.51	4.70	2.66	2.89		
D63	8.31	4.77				
D64	8.13	4.49	2.65	2.60		
D78	8.13					
D82	7.43	4.87	2.74			
D83	7.99	4.78	2.89	2.68		
D84	7.40	4.96	2.55	3.18		
D87	8.14	4.02	2.66	2.71		
D88	10.36	3.97	2.85			
D97	8.40	4.46	2.79	2.99		
D116	8.41	4.24	2.62	2.66		
E1						
E7	7.65	4.18	2.32			
E11	8.57	4.14	1.78		2.49	
E25	7.09	3.80				
E49	8.78	4.69			2.17	
E113	7.71	4.59	1.97		2.20	2.12
E121	8.88	4.83	2.04	2.19	2.34	2.42
L123	7.73	4.20	1.57			

The pK_a of the histidine residues were also tried determined.

The sample used in the determinations was approximately 14 mg BLA dissolved in 550 μl phosphate-citric acid buffer and 50 μl D_2O. This corresponded to approximately 1.6 mM BLA. pH was measured using a Sentron Argus pH meter with a Sentron Redline Standard pH probe. pH calibration was done with premade Sentron pH 4.00 and pH 7.00 buffers. No correction for deuterium isotope effect was done. The uncertainty in the pH measurement have been estimated to 0.05 units.

550 μl was transferred to a NMR tube. The pH was changed by adding a small amount of HCl or NaOH of various concentration to the NMR tube. After thorough mixing, 20 μl was removed from the NMR tube and applied to Sentron probe for pH measurement. The pH was changed 0.5 units at a time, ranging from pH 7.5 to as low as possible. At each pH value Gradshim was run and the pulse length checked, and one WATERGATE and one TOCSY/BASHD-TOCSY spectrum were recorded.

3.7.1 pK_a determinations using BASHD-TOCSY

By using BASHD-TOCSY instead of ordinary TOCSY in the pK_a determinations, one will save considerably spectrometer time. With a few exceptions all the interesting resonances are found between 2 ppm and 3 ppm. By choosing a spectral width of 2 ppm, centered at 2.5 ppm, for F1, the experiment time could in principle be one sixth of the ordinary TOCSY experiment, still keeping the same number of scans and the same spectral resolution in F1.

3.7.1.1 Acquisition

In order to get large margins in the spectra, the spectral width was set to 3.7 ppm, centered at 3.65 ppm (SPOFF1=-666.7 Hz). Most of the H^{α}-H^{N} cross peaks were then also included. Even though this made the experiment time twice as long, 5 hours and 12 minutes, it would make the assignment work easier and more reliable. The pulse program used was *dipsi2bashdxx*. Other significant acquisition parameters were:

D9	=	80 ms
D19	=	90 μs
SW(F1)	=	3.7 ppm
SW(F2)	=	14.985 ppm
TD(F1)	=	240
TD(F2)	=	1024
NS	=	56
P11	=	1.5 ms
SP1	=	17.47 dB

3.7.1.2 Processing

The spectra were Fourier transformed into 1024x1024 points using the following script:

`xf2`	FT of rows
`1 wdw no`	no window function for F1
`1 ph_mod pk`	phase correction on for F1
`1 me_mod no`	no linear prediction for F1
`xtrfp1`	FT only columns
`1 wdw QSINE`	$\cos^2$ window function (SSB=2)
`1 ph_mod no`	no phase correction
`1 me_mod LPfc`	linear prediction turned on
`xf1`	FT of columns

Before running the script the zero order phase constant for F1 was set to 0°. In order to correctly frequency shift the F1 window, the following first order phase constant was applied to F1:

$$\mathsf{phc1} = -\frac{\mathsf{SI} \cdot \mathsf{spoff} \cdot 360^\circ}{600.13 \cdot \mathsf{SW(F1)}} = 107771.5^\circ \tag{3.1}$$

A squared cosine window function was used in F2. Finally the spectra were phase corrected in F2.

3.7.1.3 Troubleshooting

The BASHD-TOCSY spectra were unsatisfactory to use to determine the pK_{a}-values. In order to try to improve the spectra, a series of BASHD-TOCSY spectra with different parameters were recorded. All the parameters were kept constant, except for the offset

of the shaped pulse and the TOCSY mixing time. Two sets of spectra were recorded, one with -130 Hz and one with -1330 Hz offset. All the spectra had a spectral width of 2.1 ppm in F1.

3.7.2 pK_a determinations using ordinary TOCSY

After unsuccessful pH-titrations using BASHD-TOCSY, regular TOCSY was tried. The TOCSY spectra were recorded with 1024 complex points, 512 increments in T1 and 64 scans. Both dimensions had a spectral width of 10.5 ppm. As for all other experiments the delay for the water suppression was 90 μs and the TOCSY mixing time was 80 ms. This gave an acquisition time of 13 hours, 2.5 times longer than the BASHD-TOCSY experiment. The spectra obtained were processed as usual (see section 3.2.3).

The titration series was stopped at pH 3.5 because of too poor quality of the TOCSY at this pH.

The TOCSY spectra for all pH values were examined in Sparky. H^N-H^β cross peaks for aspartic acid and H^N-H^β and H^N-H^γ glutamic acid were assigned by comparing with the already assigned TOCSY spectrum at pH 5.6. In cases where the H^N-H^β cross peak was difficult to assign, due to overlapping cross peaks or other factors, the H^α-H^β cross peaks assigned.

3.7.3 Curve fitting

The relevant chemical shifts were plotted against pH. Nonlinear regression analysis was used to fit the data to equation 2.29. This was done using DataFit 6.1.10 [81].

3.7.4 New attempts at low pH

To determine if it was possible to obtain usable TOCSY spectra at low pH, a new sample of BLA was made with the same concentration. Several different 1D spectra were recorded at different temperature and with different pH. The effect of adding calcium ions in order to increase the stability of the protein was also studied. The chemical shift dispersion was used as a indication of the stability of the protein. The acid denatured molten globule state will have a limited chemical shift dispersion compared with the native state. A new TOCSY spectrum was recorded at pH 3.5 and 295 K to check if it was worth the effort to record a new pH-series at this temperature.

3.7.5 Theoretical pK_a values

The theoretical pK_a values were also calculated. The calculations were preformed by Beatriz Ibarra-Molero and José M. Sánchez-Ruiz at Universidad de Granada, Spain, using the Tanford-Kirkwood model as described by Ibarra-Molero et al. [82].

Chapter 4

Results

4.1 ^{1}H-^{15}N HSQC versus FHSQC at natural abundance

A spectrum recorded with the ^{1}H-^{15}N FHSQC pulse sequence was compared with a spectrum recorded with the standard ^{1}H-^{15}N HSQC pulse sequence. All experimental conditions were identical. The 1D spectra recorded in advance of the 2D experiments were virtually identical, indicating that any differences between the two 2D spectra were due to different pulse sequences. In general the signal-to-noise ratio was a lot better for the FHSQC variant than for the regular HSQC. There was roughly a factor of 2 in difference between the signal-to-noise ratio of the two spectra for most cross peaks (see Figure 4.1). Some cross peaks, like Q65N-H^{N}, had almost completely disappeared in the noise in the regular HSQC spectrum. The line-width in the ^{15}N dimension of the cross peaks also seemed to be larger for the regular HSQC. These two factors made it harder to separate peaks in the regular HSQC spectrum (see Figure 4.2).

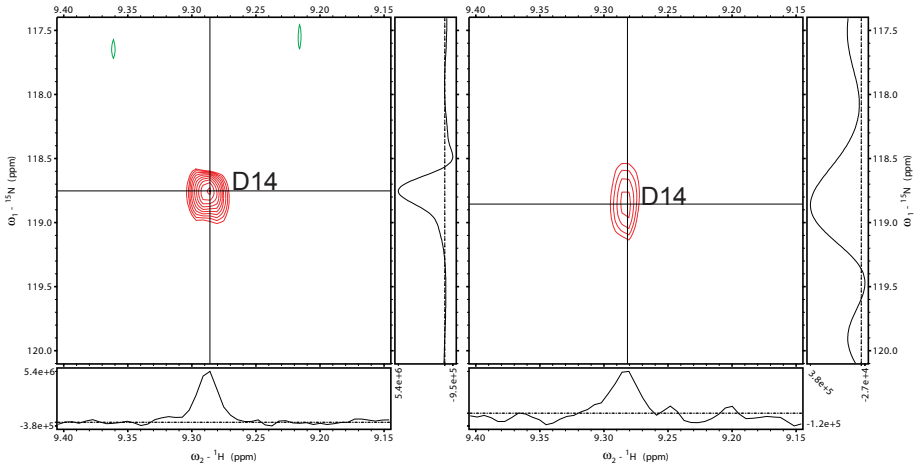


Figure 4.1: Sections of the 600 HMz ^{1}H-^{15}N FHSQC (left) and HSQC (right) spectrum of BLA at 314 K and pH 5.6. The sections show the D14N-H^{N} cross peak. The FHSQC spectrum clearly provides a better signal-to-noise ratio. It is also evident that the line-width is better for FHSQC. Both spectra are processed in an identical manner (see section 3.2.5). The lowest contour line is just above the noise level.

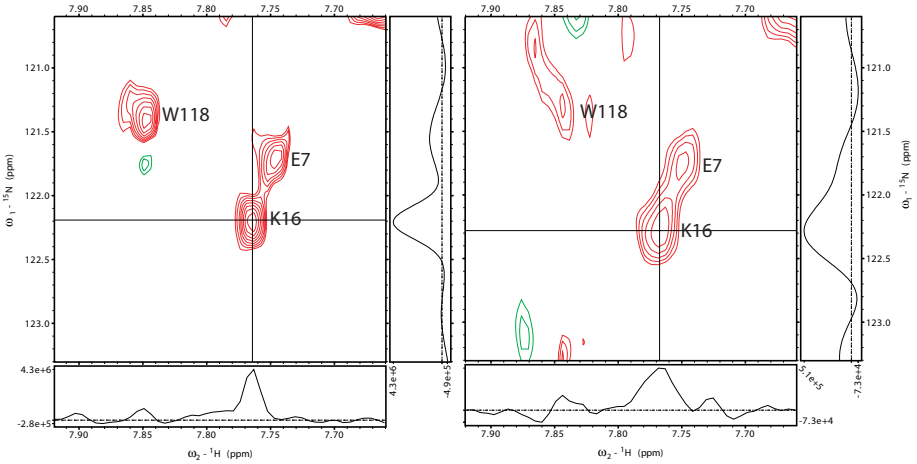


Figure 4.2: Section showing the overlap between two peaks. FHSQC to the left and HSQC to the right. K16 and E7 are harder to separate in the HSQC spectrum, and W118 has almost disappeared in the background noise.

4.2 Assignment work

Holo-BLA was assigned at pH 5.6 and 314 K using the COSY, TOCSY, NOESY and FHSQC spectra. The assignment was based upon the assignment done by Forge et al. [31] and Halskau et al. [2]. 116 of the 121 backbone amide protons and 111 of the 123 α-protons have been assigned. The spin systems of 84 of the residues have been completely, or almost completely (more than 80%), assigned.

The assignment of the β-sheet and the H3 α-helix are shown here as an illustrative example. Table 4.1 shows the assigned resonances of the backbone atoms. A complete table of the assignment can be found in the appendix.

4.2.1 Assignment of the H3 α-helix

The H^{α}-H^{N} cross peak of V92 was used as an entry point. Because this cross peak is clearly separated from other cross peaks, it was possible to assign it based on the assignment done by Forge et al. [31]. The chemical shift of the backbone nitrogen also confirmed the assignment. The H^{β} chemical shift was assigned based on the H^{α}-H^{β} cross peak in the COSY spectrum, and confirmed by the H^{N}-H^{β} cross peak in the TOCSY spectrum. The two methyl groups, $H^{\gamma 1}$ and $H^{\gamma 2}$, were assigned in the same manner, COSY cross peak with H^{β} and TOCSY cross peak with H^{β} and H^{α} (see Figure 4.3).

From V92 it was possible to do a sequential NOESY walk using the H^{N} resonances. Using only the H^{N} resonances it was possible to sequential assign all the residues from D83 to G100. The assignments were confirmed by comparing the H^{N}, H^{α} and the backbone ^{15}N resonances with the results of Forge et al. [31]. For most of the residues from L85 to L96 it was possible to find the medium range NOEs $d_{\alpha N}(i, i+2)$, $d_{\alpha N}(i, i+3)$ and $d_{\alpha N}(i, i+4)$ (see Figure 4.5). The spin system of 13 of the 18 residues from D83 to G100 were completely assigned. The three lysine residues were

Table 4.1: The assigned backbone resonances of holo-BLA at 314 K and pH 5.6. A table of all the assigned resonances can be found in the appendix.

Resid.	H^N	H^α	^{15}N	Resid.	H^N	H^α	^{15}N	Resid.	H^N	H^α	^{15}N
E1				V42	8.36	4.25	127.9	D83	7.99	4.77	
Q2				Q43	8.68	4.85	126.7	D84	7.39	4.96	117.1
L3				N44	8.72	4.99	124.0	L85	9.56	4.77	129.2
T4	7.87	4.70	107.9	N45	8.91	4.35	118.9	T86	8.51	3.73	115.4
K5	8.75	3.56	120.8	D46	8.51	4.70	117.3	D87	8.13	4.01	121.3
C6	8.57	4.66	112.7	S47	7.73	4.90	113.2	D88	10.35	3.97	
E7	7.65	4.18	121.7	T48	8.49	5.02	118.5	I89	8.67	2.84	117.2
V8	8.16	3.18	120.7	E49	8.76	4.68	124.8	M90	7.52	3.90	
F9	7.94	4.19	117.8	Y50	8.45	4.70		C91	7.88	4.80	117.8
R10	7.28	4.12	114.6	G51	9.38	4.32/4.97	112.4	V92	8.42	3.10	122.0
E11	8.57	4.14	117.1	L52	10.14	3.94	123.8	K93	7.80	3.52	117.0
L12	7.86	4.20	117.1	F53	8.75	4.54	109.3	K94	7.03	4.16	117.5
K13	6.89	3.88	119.2	Q54	8.05	3.68	117.1	I95	8.05	3.07	121.5
D14	9.19	4.57	118.7	I55	8.21	3.73	124.4	L96	9.08	3.83	121.7
L15	8.25	4.16	116.7	N56	7.48	5.93	125.8	D97	8.40	4.46	118.1
K16	7.68	3.73	122.3	N57	8.42	5.34	118.1	K98	7.96	4.53	117.9
G17	8.67	3.52	118.2	K58	8.87	4.35	119.4	V99	8.90	4.50	115.3
Y18	8.25	4.17	127.0	I59	7.47	4.29	111.2	G100	8.20	4.29	109.9
G19	8.85	3.90/3.76	118.0	W60	6.49	4.56	118.3	I101	8.30	4.42	117.6
G20	8.45	4.82/3.32	127.2	C61	7.91	5.82	111.8	N102	8.27	4.64	117.2
V21	7.05	3.81	122.5	K62	8.19	4.94	122.6	Y103	7.60	3.86	120.3
S22	8.14	4.60	122.5	D63	8.33	4.77	127.8	W104	8.09	4.42	117.2
L23	9.52	4.54	120.1	D64	8.13	4.45	114.9	L105			
P24				Q65	8.24	3.85	119.1	A106	8.75	4.06	116.6
E25	7.09	3.80	113.3	N66	8.07	5.37	114.3	H107	7.82	4.83	115.2
W26	8.14	4.43	123.2	P67				K108			
V27	8.77	3.55	123.7	H68	7.48	3.31		A109			
C28	7.68	2.02	120.2	S69	7.05	4.40	113.3	L110	7.62	4.61	113.8
T29	8.34	3.58	117.8	S70	8.08	4.21	122.0	C111	7.32	5.13	116.6
T30	8.69	4.12	113.4	N71	7.97	3.66	117.0	S112	8.08	4.55	111.3
F31	7.77	3.80	125.2	I72	8.71	3.63	120.0	E113	7.71	4.59	118.3
H32	7.36	4.10	116.0	C73	9.45	4.29	113.5	K114	8.80	4.14	118.4
T33	7.97			N74	7.96	4.17	120.4	L115	8.44	4.53	119.1
S34				I75	9.31	4.35	121.4	D116	8.41	4.24	117.5
G35	7.90	3.27/3.65		S76	8.71	5.00	119.9	Q117	7.91	4.06	115.9
Y36	7.12	3.90	106.1	C77	8.45		123.3	W118	7.77	4.58	121.2
D37	7.08	5.14	118.7	D78	8.13		115.4	L119	6.93	4.28	115.3
T38	8.49	4.07	115.4	K79	7.56	4.12	120.5	C120	6.63	4.53	118.5
Q39	8.31	4.65	115.2	F80	8.17	5.63	121.7	E121	8.88	4.84	123.9
A40	6.97	4.03	122.7	L81	7.00	5.73	113.7	K122	7.92	4.33	121.0
I41	7.97	4.88		D82	7.44	4.87	118.6	L123	7.73	4.20	129.6

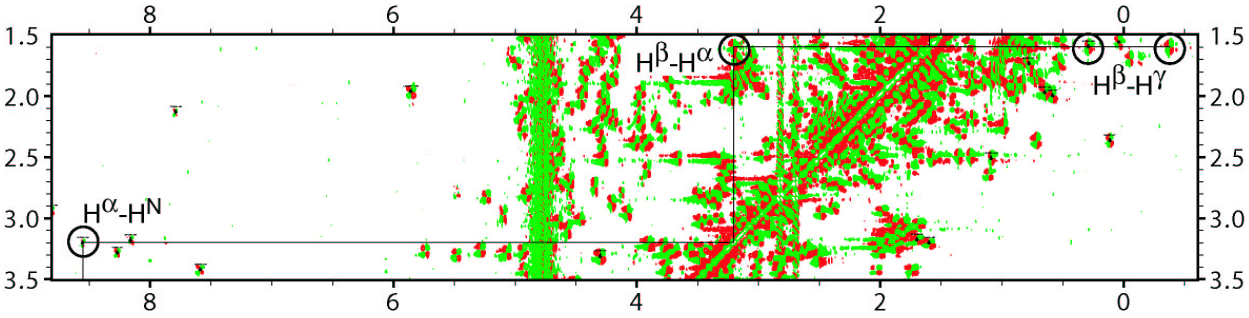


Figure 4.3: 600 MHz DQF-COSY spectrum of BLA showing the spin system of V92. Because of large chemical shift dispersions this residue was relatively easy to assign. Most residues had cross peaks in the very crowded region close to the diagonal. This made them virtually impossible to assign fully.

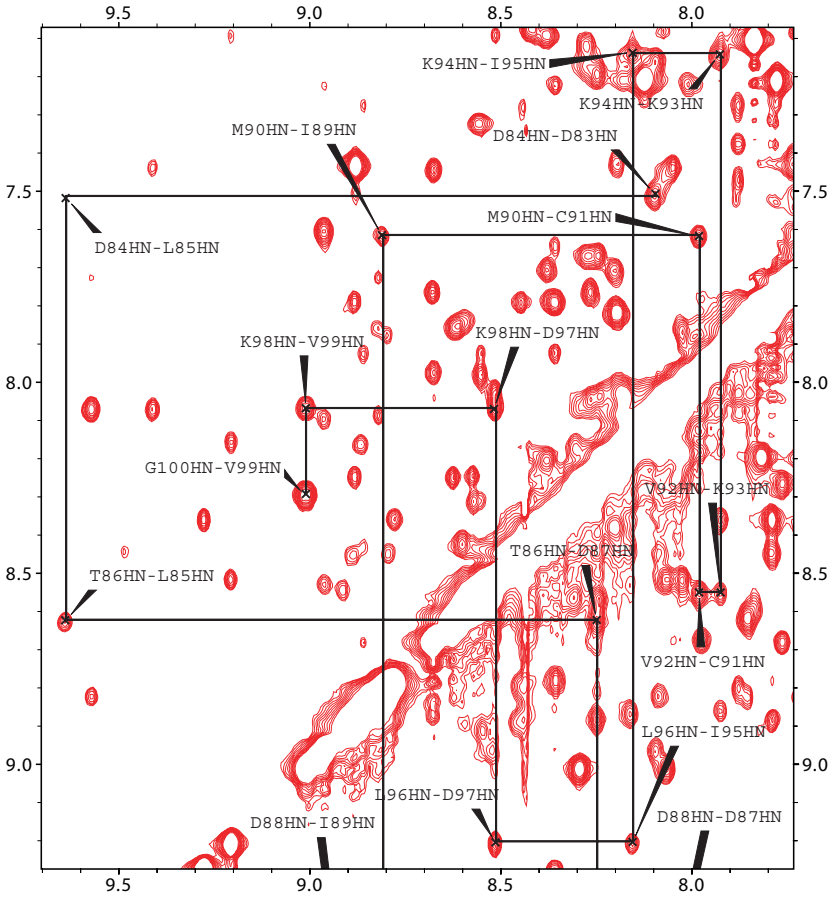


Figure 4.4: A section of the NOESY spectrum (150 ms mixing time) of BLA showing the sequential assignment of the residues from D83 to G100. Two of the cross peaks ($D88H^{N}$-$I89H^{N}$ and $D88H^{N}$-$D87H^{N}$ are located outside of the displayed region.

the hardest to assign, none of these were completely assigned. L85 and I89 were not completely assigned either.

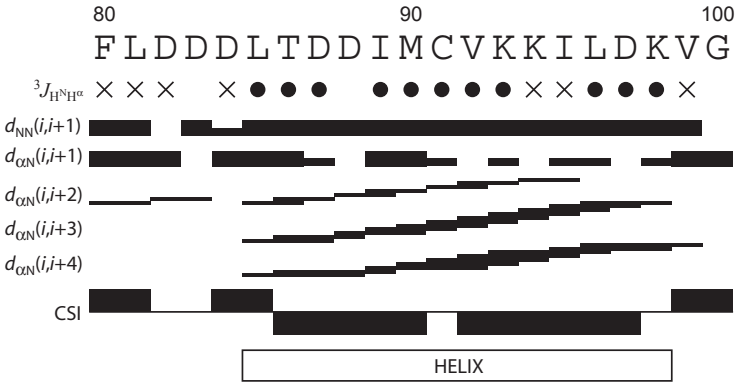


Figure 4.5: Graphical representation of the sequential NOEs for the residues from F80 to G100. Thick bars indicate intense NOEs. The medium-range NOEs which indicate secondary structure are also shown. Residues with $^3J_{\mathrm{H^N H^\alpha}}$<6.0 Hz are represented by filled circles, and those with 6.0 Hz<$^3J_{\mathrm{H^N H^\alpha}}$<8.0 Hz by crosses (see section 4.3). The Chemical Shift Indices (see section 4.4) are shown at the bottom.

Residue I95 have some strange chemical shifts. One of the methylene protons has a chemical shift of -2.41 ppm. The other one resonates at 0.60 ppm. The normal chemical shift for the methylene protons in isoleucine is $1.0-1.5$ ppm. The very upfield methylene proton also has NOEs to the four aromatic protons on the six-membered ring of W60.

The NOE correlations of the methylene proton with the aromatic protons are almost equal (volume of cross peaks in arbitrary scale: $1.0{\cdot}10^7$, $1.3{\cdot}10^7$, $1.3{\cdot}10^7$ and $1.4{\cdot}10^7$). Using the proton distances on the aromatic ring, and their corresponding NOEs, as a calibration factor, the distance to the methylene proton can be calculated:

$$d_{\mathrm{obs}} = \sqrt[6]{\frac{C}{V_{\mathrm{obs}}}} = \sqrt[6]{\frac{V_{\mathrm{ref}} d_{\mathrm{ref}}^6}{V_{\mathrm{obs}}}} = \sqrt[6]{\frac{1.0 \cdot 10^8 \cdot (2.4\ )^6}{1.3 \cdot 10^7}} \approx 3.4 \qquad (4.1)$$

In this calculation the volume of the four cross peaks have been assumed equal ($1.3 \cdot 10^7$). A simple geometric calculation then shows that the methylene proton is situated about 2.4 Å above the center of the six-membered ring, slightly to one of the sides.

4.2.2 Assignment of the β-sheet

T48H$^\alpha$-H$^\mathrm{N}$ was used as an entry point assignment of the β-sheet region. This cross peak had chemical shifts close to the values in the literature, and was well separated from other cross peaks. The chemical shifts of the side chain (H$^\beta$=3.89 ppm and H$^{\gamma 2}$=0.64 ppm) confirmed that the residue was threonine. Using only $d_{\alpha\mathrm{N}}(i, i+1)$ connectivities it was possible to assign all the residues from A40 to K58. The only problem was connecting Y50 and G51. Because of distortions from the water signal, it was not possible to observe a Y50H$^\alpha$-G51H$^\mathrm{N}$ cross peak. Y50H$^\mathrm{N}$ had a weak NOE correlation to a signal at 9.37 ppm. This signal had a spin system which proved that it was a glycine residue, most likely G51.

The residues from A40 to K58 also had an extensive network for long-range backbone NOEs ($d_{\alpha\alpha}(i,j)$, $d_{\alpha\mathrm{N}}(i,j)$ and $d_{\mathrm{NN}}(i,j)$). Figure 4.6 shows the sequential NOEs confirming the assignment and the long-range NOEs connecting the different parts of the chain. The spectrum showing the NOE correlations is not shown.

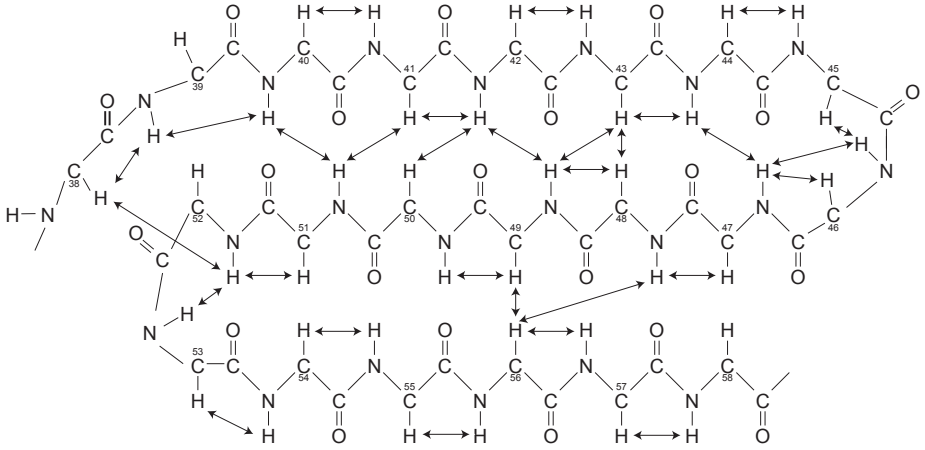


Figure 4.6: The sequential, medium-range and long-range NOEs connecting the residues in the β-sheet.

4.3 ACME

The software package ACME was used to measure ${}^{3}J_{\mathrm{H}^{\alpha}\mathrm{H}^{\mathrm{N}}}$ the coupling constants from the phase sensitive COSY spectrum. The measurement consisted of two steps. First a diagonal-only spectrum was used to estimate the intrinsic intensity, then the coupling constants were measured by restraining the intensity of the multiplets.

4.3.1 Estimation of intensity

There was a large variation in the measured intensities. G51H^{N} and N57H$^{\delta 2}$ constituted the lower extreme, with an intrinsic intensity of 366 and 370 respectively (arbitrary scale). This was not unexpected considering that no cross peak could be found for any of these resonances in the COSY spectra. The two next resonances with low intensities, L85H^{N} and L23H^{N} with an intensity of 617 and 556 respectively, had no COSY cross peaks either. In the other end of the scale, Y103H$^{\delta}$ was found, with an intensity of 1788. The rest of the resonances had intensities ranging from 750 to 1200, with the range 900-1000 as a mode (see Table 4.2). The 1000 was chosen as the intensity to use in the measurement of the coupling constants.

4.3.2 Measurement of coupling constants

Many of the cross peaks found in the DQF-COSY spectrum were impossible to find in the COSY spectrum used with ACME. This mostly applied to cross peaks close to the water resonance. Of the 93 assigned H$^{\alpha}$-H$^{\mathrm{N}}$ cross peaks in the gs-DQF-COSY spectrum, 76 could be found in the spectrum used with ACME. All but two of the 17

Table 4.2: Intensities of resonances used to estimate the intrinsic intensity.

Resonance	Intensity
N57H$^{\delta 2}$	370
W60H$^{\epsilon 1}$	762
L85H^{N}	617
L23H^{N}	556
C73H^{N}	836
G51H^{N}	366
I75H^{N}	950
D14H^{N}	961
L96H^{N}	1143
6.37 ppm	930
N56H$^{\alpha}$	1183
Y103H$^{\delta}$	1788

missing peaks were located within 0.15 ppm from the water resonance. Only three peaks were visible within the same range.

Of the remaining 76 cross peaks, 12 were found to be very weak. Of these 12 cross peaks, 10 belonged to residues that, according to the x-ray structure [4], are clearly solvent accessible.

All the coupling constants were found to be in the range between 2.2 Hz and 9.4 Hz (see Table 4.3). And all but one of the cross peaks gave a sensible result. When measuring the coupling constant for N56H$^{\alpha}$-H^{N}, the resulting *Y Width* was negative. This happened no matter what the starting values were. The *X Width* of this cross peak is 16.1 Hz and the *Y Width* is -1.3 Hz. The active coupling constant is 3.9 Hz, while the passive Y coupling constant 13.9 Hz (see Figure 4.7).

ACME also aided in the discovery of new cross peaks. In the difference spectrum, when fitting L15H$^{\alpha}$-H^{N}, another cross peak, partly overlapping with L15H$^{\alpha}$-H^{N}, was revealed. This was the Y18H$^{\alpha}$-H^{N} cross peak. Another similar case was found in the bunch of signals where F9H$^{\alpha}$-H^{N} and N74H$^{\alpha}$-H^{N} are located. The shape of the broad TOCSY peak for F9 and N74 indicated that there was at least one other cross peak hidden there somewhere. This hidden, unknown cross peak was found when fitting the H$^{\alpha}$-H^{N} COSY peak of F9, S70, N74 and Q117 (see Figure 4.8).

The effect of decreasing the intensity factor was an increase in the coupling constant. Changing the intensity factor from 1000 to 900, caused the coupling constant of K94H$^{\alpha}$-H^{N} to increase with 0.29 Hz to 6.54 Hz. The coupling constant of D87H$^{\alpha}$-H^{N} increased with 0.22 Hz to 4.63 Hz. Figure 4.9 shows a plot of coupling constant as a function of intrinsic intensity for the two residues.

4.4 CSI

The chemical shift index was calculated in order to introduce additional torsion angle restraints into the structural calculations.

A graphical view of the results from the secondary structure determination using CSI can be found in Figure 4.10. Residues indicated with an index +0.1 are residues that have no assigned H$^{\alpha}$ chemical shift. These were treated as if they had index 0.

There is clearly a large α-helix at the beginning of the sequence, between residue 5 and 20. This helix is disrupted by a region with random coil before the helix continue

Table 4.3: Table of the $^3J_{\mathrm{H}^\alpha\mathrm{H}^\mathrm{N}}$ measured with ACME

Residue	$^3J_{\mathrm{H}^\alpha\mathrm{H}^\mathrm{N}}$	Residue	$^3J_{\mathrm{H}^\alpha\mathrm{H}^\mathrm{N}}$	Residue	$^3J_{\mathrm{H}^\alpha\mathrm{H}^\mathrm{N}}$
7E	6.5	56N	3.9	90M	4.2
8V	5.1	57N	6.3	91C	3.6
9F	6.8	58K	7.1	92V	4.9
10R	4.8	59I	8.6	93K	4.6
11E	4.5	61C	8.1	94K	6.2
12L	7.7	62K	7.0	95I	6.2
15L	7.9	63D	3.3	96L	3.7
16K	2.2	65Q	3.8	97D	3.3
18Y	3.3	66N	9.2	98K	4.6
25E	7.6	70S	3.7	99V	6.5
26W	5.1	71N	6.0	101I	2.8
27V	4.4	72I	3.8	103Y	4.2
32H	5.8	73C	7.0	104W	6.6
36Y	6.0	74N	7.3	106A	3.9
37D	6.5	75I	6.3	107H	4.5
38T	3.6	76S	6.1	111C	7.7
40A	2.6	79K	5.7	112S	4.2
41I	6.2	80F	7.1	114K	6.2
42V	9.4	81L	8.2	115L	3.2
43Q	6.5	82D	6.6	116D	3.8
44N	7.0	84D	7.8	117Q	5.1
47S	6.4	85L	3.1	119L	6.5
48T	7.9	86T	3.4	121E	5.0
53F	4.2	87D	4.4	122K	7.5
54Q	5.8	89I	4.0	123L	8.1
55I	6.0				

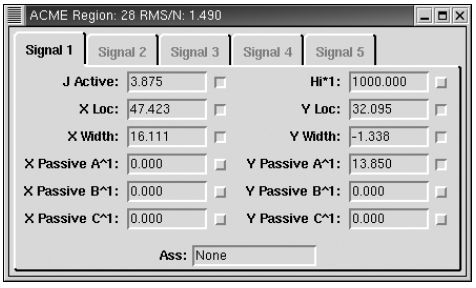


Figure 4.7: The parameter window with the result of the fitting of peak N56H$^\alpha$-H$^\mathrm{N}$.

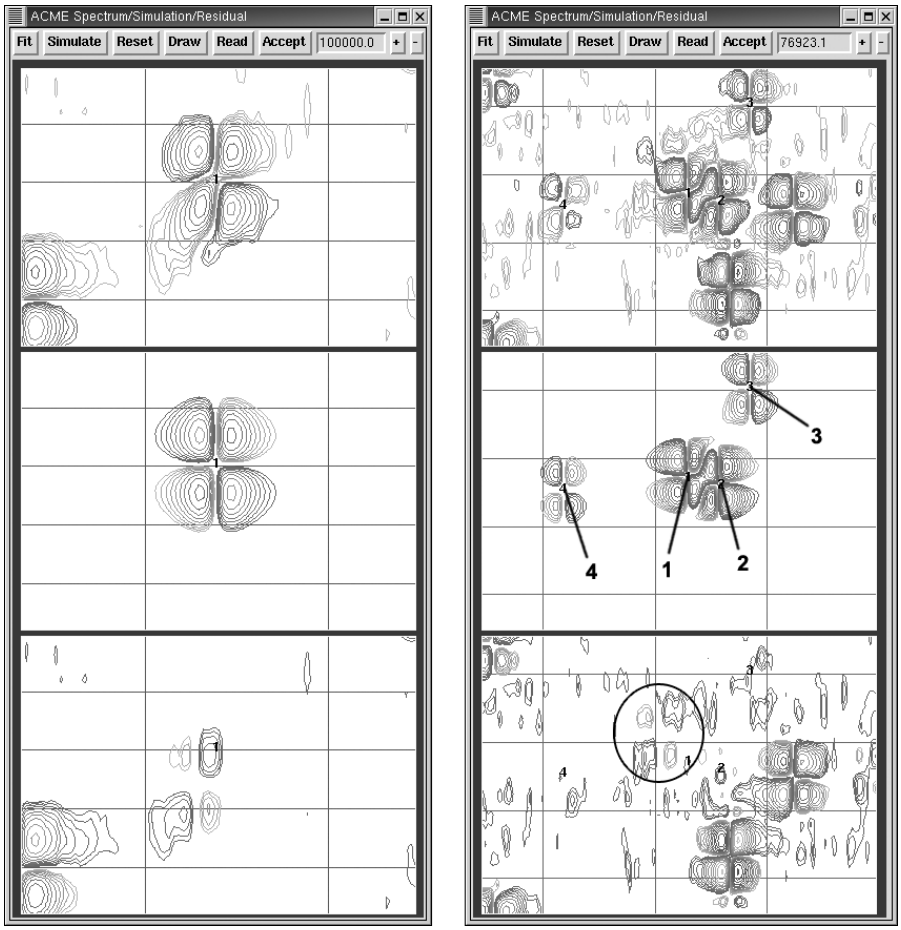


Figure 4.8: Expreimental data (top), fitted data (middle) and difference spectrum (bottom) of ACME. The left windows show the fitting of L15. The difference spectrum reveals a new peak partly overlapping with L15. The right windows show the fitting of N74 (1), F9 (2), Q117 (3) and S70 (4). An unknown signal can be seen above to the left of N74.

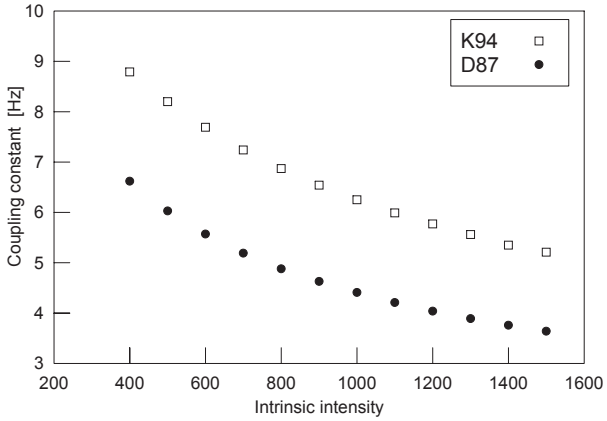


Figure 4.9: The measured coupling constant plotted as a function of intrinsic intensity for the two cross peaks D87H$^{\alpha}$-H$^{\mathrm{N}}$ and K94H$^{\alpha}$-H$^{\mathrm{N}}$

from residue 25 to 32. The next three residues have no assigned H$^{\alpha}$ resonances. It is therefore not possible to say anything about the structure. A β-sheet between residue 41 and 44 is the next ordered structure indicated by CSI. Disrupted by two residues it continues to G51. This β-sheet is followed by four residues with α-helix structure and four with β-sheet structure.

The Chemical shift index predicts another α-helix from S69 to N74. Then following a region with unknown structure there is a large α-helix found between T86 and D97. The rest of the sequence have only smaller regions with ordered structure. There are small β-sheets, V99-I101 and L110-E113, and two α-helices, N102-A106 and D116-C120.

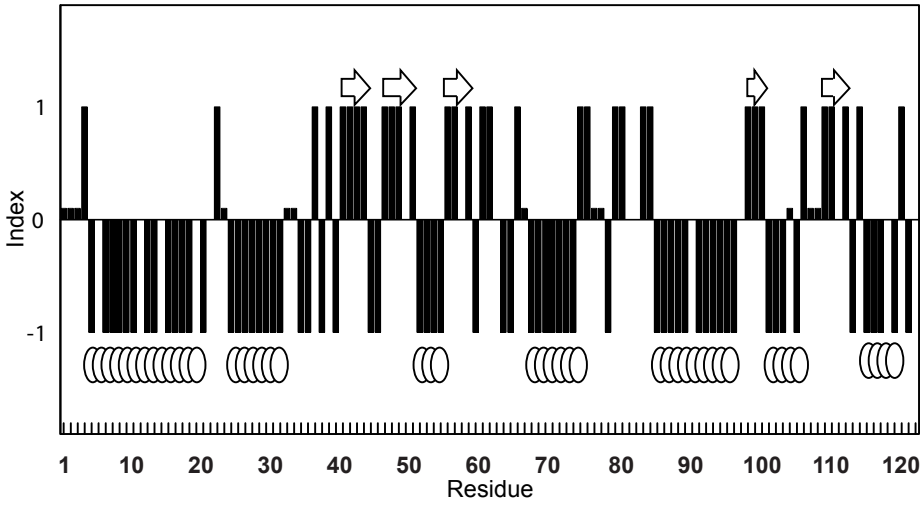


Figure 4.10: Results of the Chemical Shift Index, CSI. Local densities of "-1" indicate α-helix, and densities of "+1" indicate β-sheet. The determined secondary structures are shown (circles=α-helices and arrows=β-sheets).

4.5 Structural calculations

The structural calculations were preformed with the software ARIA and CNS. During the calculations unlikely restraints are removed by ARIA. These are restraints that may arise because of spectral artifacts, incorrect resonance assignments and missing proton assignments. The structures of the final iteration were calculated using 2275 NOE restraints. 803 of these restraints were ambiguous. Table 4.4 shows how these restraints were distributed among intraresidual, sequential, medium range and long range NOEs.

Table 4.4: NOE restraints used to calculate the structures of the last iteration. The numbers in parenthesis are the numbers of ambigiuous restraints.

Intraresidual	504.3	(81.4)
Sequential	476.7	(137.7)
Medium-range	304.0	(124.0)
Long-range	993.9	(459.9)

The lowest energy structure, of the 20 structures calculated in the last iteration, violated 331 of the NOE restraints (222 unambiguous and 109 ambiguous). 44 of the violations were greater than 0.5 Å. The root mean square deviation (RMSD) was 0.273 Å.

The lowest energy structure also violated 63 of the 130 dihedral restraints derived from CSI (RMSD = 16.6°) and 25 of the 44 coupling constant restraints (RMSD = 1.82 Hz).

A superimposition of the 7 structures with lowest energy is shown in Figure 4.11. Different colouring are used to indicate the precision of the different parts of the structure. The N-terminal and C-terminal of the protein have a very high RMDS. The highest RMSDs are found for the N-terminal residue, E1, with a RMSD of 8.3 Å. The most unstructured region is the C-terminal. All the residues from H107 to L123 have a RMSD for the backbone atoms which is higher than 2.0 Å. There are also three regions, in addition to the N- and C-terminal regions, which have a large RMSDs (RMSD>2.0 Å): 17-20, 42-47 and 65-67. The average RMSD for the backbone atoms of all residues is 1.6 Å. When including the side chains the average RMSD is 2.1 Å. The average RMSDs, when excluding the poorly defined terminals (residue 1-6 and 107-123), is 1.1 Å and 1.5 Å for the backbone atoms and all atoms respectively. The average RMSDs for the well defined H3 α-helix is 0.4 Å and 0.7 Å. The RMSDs for the backbone atoms and all the atoms in each residue are plotted in Figure 4.12.

4.6 pK_a determinations

A series of TOCSY spectra were recorded at different pH values in order to determine the pK_a values of the aspartic and glutamic acid residues by plotting their chemical shifts as a function of pH.

4.6.1 pK_a determinations using BASHD-TOCSY

The BASHD-TOCSY variant was first attempted. Compared with using ordinary TOCSY this would lead to a considerable time gain. However, the BASHD-TOCSY spectra were not of satisfactory quality.

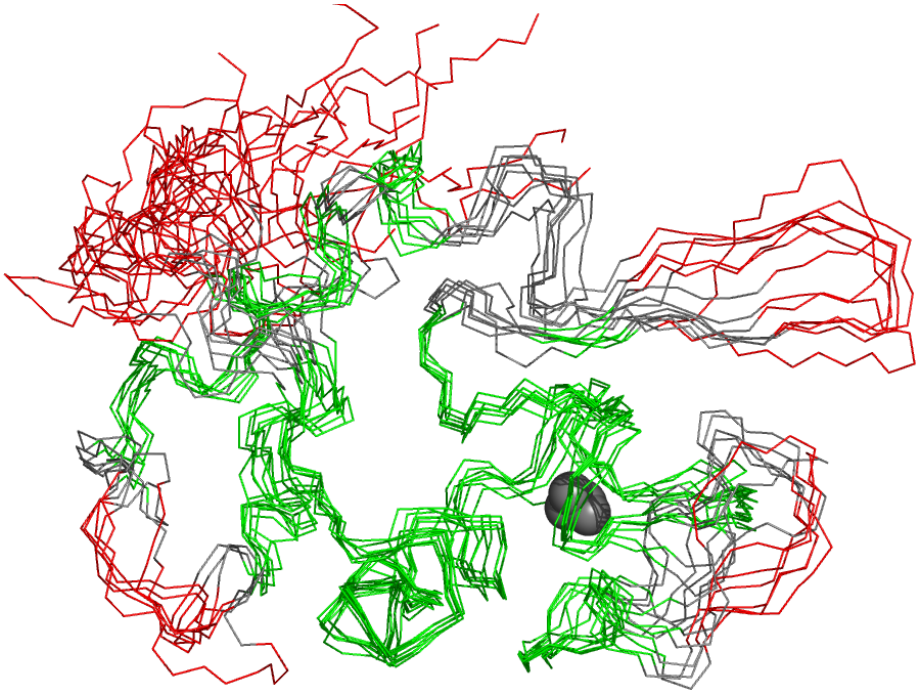


Figure 4.11: Superimposition of the 7 lowest energy structures of the 20 calculated structures. The protein is viewed from approximately the same angle as figure 1.11. Residues with backbone atoms with RMSDs less than 1.0 Å are coloured green. Red indicates residues with backbone RMSD larger than 2.0 Å.

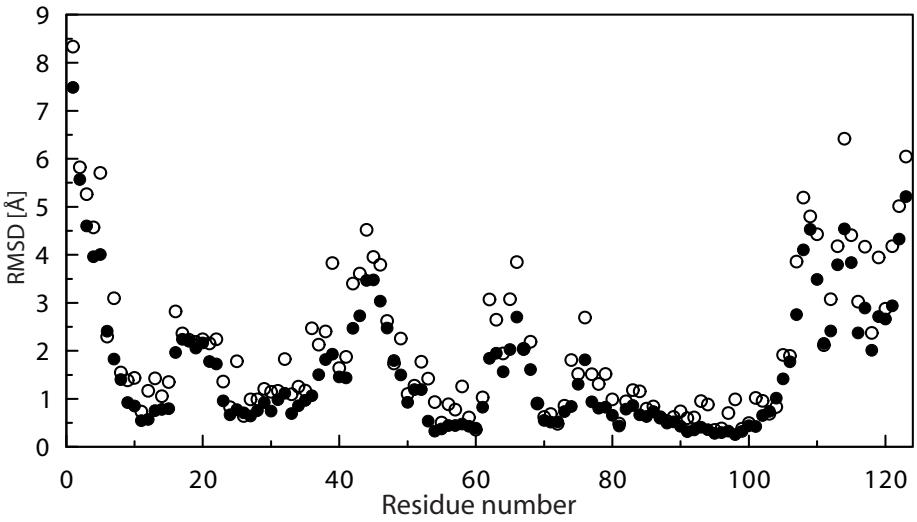


Figure 4.12: Plot of the backbone RMSDs (●) and the mean RMSD of all atoms in residue each residue (○).

Close to the water resonance in the F1 dimension the quality of the BASHD-TOCSY spectra was good. The cross peaks were of almost the same quality as the cross peaks in an ordinary TOCSY spectrum. Almost all the cross peaks found in the TOCSY spectrum used in the assignment work could be located in the BASHD-TOCSY spectra. In addition there were a few new cross peaks that had been folded in. As the chemical shifts deviated from *O1P*, the peak intensities gradually got worse compared with an ordinary TOCSY spectrum. Around 2 ppm, only the very strong peaks were visible. By using these BASHD-TOCSY spectra, it would only be possible to determine a very limited number of the pK_a-values, if any at all.

A series of BASHD-TOCSY spectra were recorded with different acquisition parameters in order to find parameters that would give spectra of satisfactory quality. With an offset of -130 Hz, the spectra were almost of the same quality as the ordinary TOCSY spectrum. The intensities were in general a little weaker, but not to a great extent. A small number of cross peaks had completely vanished. The vanished cross peaks seem mainly to be the H^{β}-H^{N} cross peaks. In the range that is covered by the BASHD-TOCSY spectra, 3.5 ppm to 5.6 ppm, eleven H^{β}-H^{N} cross peaks were found in the TOCSY spectrum from the assignment work. Only one of these cross peaks was found in the BASHD-TOCSY spectra, S47H^{β}-H^{N}, and that cross peak was very weak. Ten of the H^{α}-H^{N} cross peaks had also disappeared.

Only minor differences were found between the different TOCSY mixing times. Where differences were found, the spectrum with 40 ms mixing time was found to be better than the one with 80 ms mixing time.

In the spectra with -1330 Hz offset, almost no cross peaks to the amide protons were found. The exceptions were a few cross peaks that were really strong in the ordinary TOCSY spectrum, like L123$H^{\beta 2/\beta 3}$-H^{N} and K122$H^{\beta 2/\beta 3}$-H^{N}. In the other parts of the spectra, there were many more cross peaks to be found, especially on the opposite side of the diagonal. Below 2 ppm most cross peaks in the TOCSY spectrum could also be found in the BASHD-TOCSY spectra, regardless of the TOCSY mixing time. Because of extensive overlapping the exact number of missing peaks was hard to decide. Of the eight cross peaks found in the area (0.5, -0.5)x(1.5, 3.5) in the TOCSY spectrum, only one was missing in the BASHD-TOCSY spectra.

4.6.2 pK_a determinations using ordinary TOCSY

As the pH deviated from pH 5.6, at which the assignments originally had been done, it became increasingly more difficult to assign the cross peaks. At pH 7.53 and 3.70 some resonances were impossible to assign, either because of overlapping with other cross peaks or simply because they disappeared. For one of the D97 β protons, it was not possible to find any cross peak below pH 4.8. One example of how the spectra looked like is shown in Figure 4.13. This shows how the D46H^{β}-H^{N} cross peaks move when the pH is changed. Notice that the signal has completely disappeared at pH 7.53. D46H^{β}-H^{N} was one of the easier cross peaks to assign.

At pH $\geq$ 6.38 E7H^{β}-H^{N} had only one cross peak. At pH 5.88 the β protons seemed to have slightly different chemical shifts. When the pH was decreased further, there were clearly two different signals. One of them was partly overlapping with E113H^{β}-H^{N}. This made it difficult to determine the chemical shift accurately enough. On this basis, only the β proton not overlapping was used in the pK_a determination (see Figure 4.14).

D87H^{β}-H^{N} was a similar case. At high pH the two β-protons had different chemical shifts. But as the pH decreased the two protons got the same chemical shift. In the

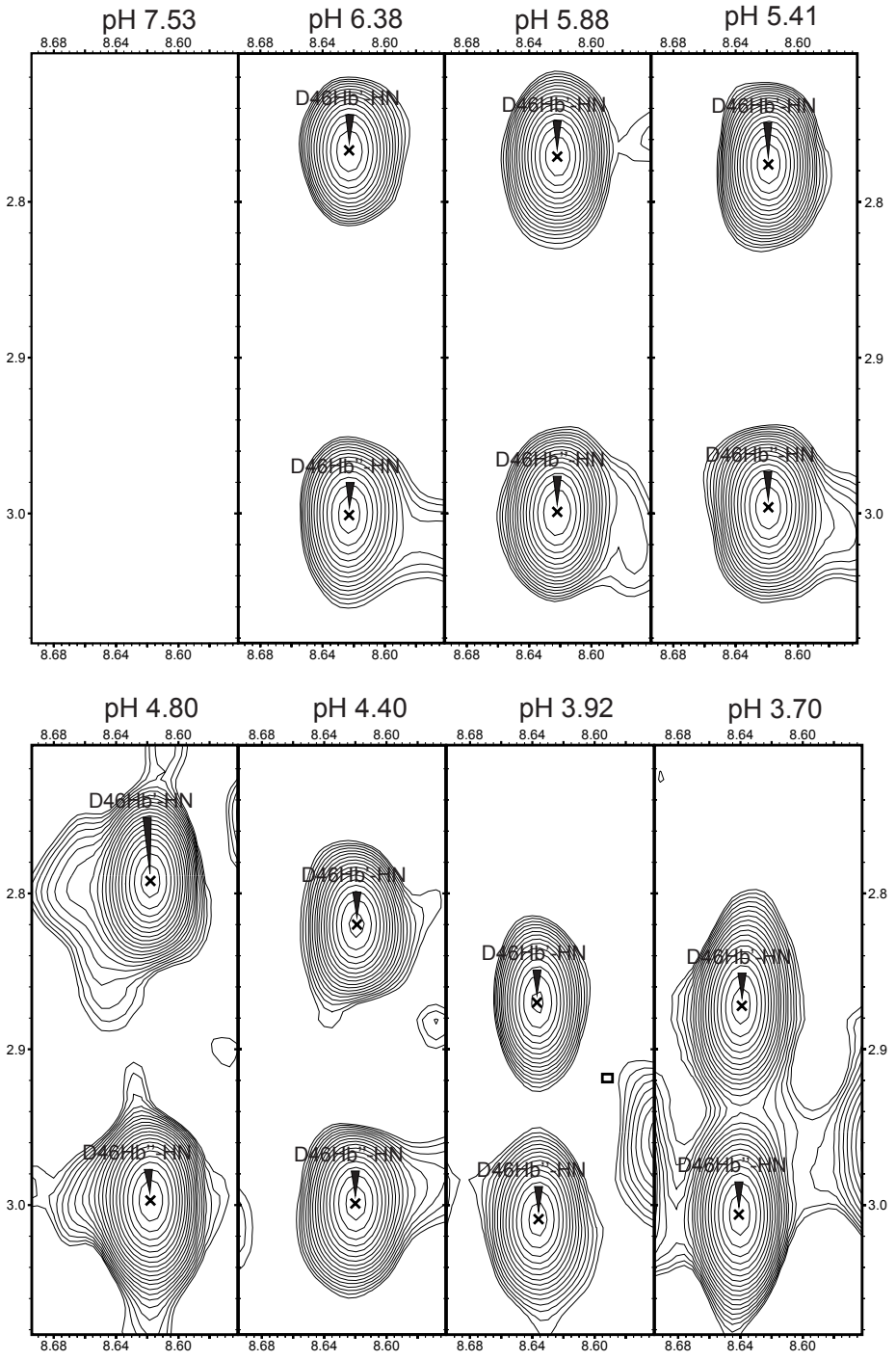


Figure 4.13: Dependency of the D46H$^{\beta}$-H^{N} cross peaks on pH. The cross peaks have disappeared at pH 7.53. One of the H$^{\beta}$ shifts is possible to find by looking in the H$^{\alpha}$-H$^{\beta}$ region.

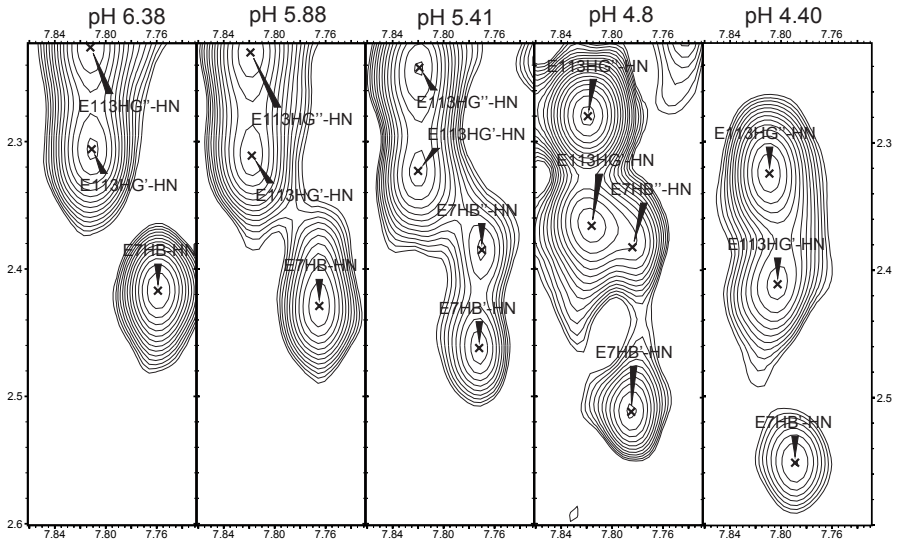


Figure 4.14: Figure showing the partial overlap of E7H$^{\beta 2/\beta 3}$-H$^{\mathrm{N}}$ and E113H$^{\beta 2/\beta 3}$-H$^{\mathrm{N}}$. At high pH the two β-protons of E7 have the same chemical shifts.

pK_a determination, the average chemical shift value was used where there were two signals.

Close examination of the D87H$^{\beta}$-H$^{\mathrm{N}}$ cross peaks in TOCSY spectra revealed that there were two sets of cross peaks with sigmoidal behavior. They both had the same H$^{\mathrm{N}}$ shift (around 8.14 ppm) at high pH, but slightly different shift at lower pH-values. Based on the fact that there was a NOE correlation with D63H$^{\alpha}$, a reasonable conclusion was that the second set of cross peaks was D64H$^{\beta 2/\beta 3}$-H$^{\mathrm{N}}$. This also made it possible to identify the D64H$^{\alpha}$ resonance. The D64H$^{\alpha}$-H$^{\mathrm{N}}$ cross peak was overlapping with W26H$^{\alpha}$-H$^{\mathrm{N}}$ and W104H$^{\alpha}$-H$^{\mathrm{N}}$. During the pH titration it became clear that this broad peak constituted of at least three different cross peaks.

When plotting the titration curves a little typing error was revealed. The residue assigned as D66 did not demonstrate a sigmoidal behavior. One of the β protons increased linearly with pH, the other decreased linearly. A look at the protein sequence showed that D66 was in fact not aspartic acid, but asparagine. The same was the case for E39. A few other residues too did not show a sigmoidal behavior. However, no evidence was found supporting the fact that these had been assigned wrongfully.

4.6.3 Curve fitting

When plotting the titration curves, it became clear that something was wrong with the spectrum at pH 3.70. All the titration curves indicated that the pH measured was wrong. The actual pH was probably closer to 4.0. Because of this, the spectrum at pH 3.70 was ignored in the pK_a determinations.

The titration curves, together with the best fit from the regression analysis, are plotted in Figures 4.15 and 4.16. No curve is shown for H107. The values from the regression analysis can be found in Table 4.5. The uncertainties of the pH measurement were not taken into account when the pK_a values were calculated.

Many of the residues show no sign of a sigmoidal behavior. This is visualized

Table 4.5: Results from the regression analysis for the aspartic and glutamic acid residues. *n* is the Hill coefficient, while δHA and δA are the chemical shifts when fully protonated and fully deprotonated respectively. The correlation coefficients, R^2, are also shown.

Resid.	Res.	pK_a	*n*	δA	δHA	R^2
D14	$H^{\beta'}$	4.3±0.2	4±8	2.575±0.002	2.613±0.008	0.9766
	$H^{\beta''}$	0±30	0.3±0.2	2.63±0.01	0±45	0.9934
D37	$H^{\beta'}$	4.2±0.1	0.86±0.08	2.594±0.002	2.89±0.03	0.9993
	$H^{\beta''}$	4.3±0.1	1.2±0.3	3.266±0.001	3.35±0.02	0.9968
D46	$H^{\beta'}$	3.8±0.3	0.9±0.1	2.767±0.002	3.01±0.09	0.9994
	$H^{\beta''}$	$0 \pm 1 \cdot 10^7$	$0 \pm 4 \cdot 10^9$	2.9985±0.0009	$0 \pm 3 \cdot 10^5$	0.8591
D63	H^N	4.5±0.2	2±1	8.431±0.007	8.52±0.02	0.9410
D64	$H^{\beta'}$	3.9±0.3	0.7±0.1	2.705±0.002	2.96±0.07	0.9995
	$H^{\beta''}$	4.20±0.06	0.92±0.05	2.8040±0.0005	2.940±0.007	0.9998
D78	No usable resonances					
D82	H^{β}	$0 \pm 1 \cdot 10^4$	$0 \pm 3 \cdot 10^5$	2.839±0.003	$0 \pm 9 \cdot 10^{11}$	0.4664
D83	$H^{\beta'}$	5.0±0.2	0.47±0.08	2.932±0.007	3.11±0.02	0.9984
	$H^{\beta''}$	4.4±0.3	2±3	2.797±0.008	2.89±0.04	0.8908
D84	$H^{\beta'}$	4.1±0.2	1.1±0.2	2.661±0.001	2.76±0.02	0.9979
	$H^{\beta''}$	4.4±0.1	2±1	3.288±0.001	3.322±0.006	0.9747
D87	H^{β}	4.40±0.08	1.3±0.2	2.792±0.001	2.90±0.01	0.9981
D88	H^{β}	$0 \pm 1 \cdot 10^3$	1±2	2.967±0.008	$0 \pm 4 \cdot 10^3$	0.9216
D97	$H^{\beta'}$	5±1	2±7	2.904±0.004	2.91±0.01	0.6456
	$H^{\beta''}$	0±30	0.6±0.6	3.101±0.003	0±30	0.9739
D116	H^{β}	4.38±0.08	2.1±0.7	2.744±0.002	2.809±0.008	0.9909
E7	H^{β}	4.9±0.1	1.1±0.2	2.421±0.006	2.59±0.01	0.9937
E11	H^{β}	4.7±0.3	0.5±0.1	1.858±0.006	2.03±0.03	0.9973
E11	H^{γ}	$0 \pm 5 \cdot 10^7$	$0 \pm 4 \cdot 10^7$	2.60±0.02	$0 \pm 1 \cdot 10^{11}$	0.0977
E25	H^N	4.9±0.3	0.6±0.2	7.16±0.01	7.30±0.02	0.9932
E49	H^{γ}	4±1	0.6±0.2	2.252±0.004	2.6±0.3	0.9980
E113	$H^{\gamma'}$	4.3±0.1	1.0±0.1	2.307±0.002	2.55±0.03	0.9990
	$H^{\gamma''}$	3.9±0.1	0.83±0.04	2.223±0.001	2.58±0.04	0.9998
E121	$H^{\gamma'}$	Regression analysis failed				
	$H^{\gamma''}$	4.3±0.9	1±3	2.535±0.003	2.56±0.03	0.7896
H32	$H^{\delta 2}$	5.79±0.04	1.4±0.2	8.15±0.02	8.75±0.01	0.9977
H107	H^{N}	6.02±0.05	1.0±0.1	7.71±0.01	8.04±0.01	0.9980
L123	H^{N}	4.14±0.05	0.99±0.04	7.8407±0.0005	8.04±0.01	0.9999

both in the titration curves, and in the crazy values and large standard errors in the regression analysis (see Table 4.5). Obviously these will give unreliable pK_a-values. The regression analysis for E121H$^{\gamma'}$ failed.

Ideally a carboxyl group not affected by other titratable groups should have a Hill coefficient, *n*, around 1. D83H$^{\beta'}$ has a very low Hill coefficient, only 0.47. Other titration curves which have a low Hill coefficient are E11H$^{\beta}$, E25H$^{\mathrm{N}}$ and E49H$^{\gamma}$.

Because the last point in the titration had a pH as high as 3.90, non of the resonances have reached their fully-protonated chemical shift. This makes all the pK_a-values somewhat uncertain. For some of the resonances there seems to be other factors influencing the chemical shift too. Both D64H$^{\mathrm{N}}$ and D83-H$^{\beta''}$ have an increase in the chemical shift above pH 6, where it should be stable.

4.6.4 pH-stability

The volumes of the cross peaks gradually decreased as the pH was decreased. The relative volumes were impossible to determine because of the lack of cross peaks with constant volume to compare with. It was easier to measure the relative intensities in the WATERGATE spectra. These intensities were compared with the intensities of the signals from citric acid, which were assumed to be constant. Figure 4.17 shows the relative intensities for the methyl groups of V92 and one of the methyl groups of I95 as a function of pH.

4.6.5 Theoretical pK_a values

The results of the theoretical pK_a calculations are found in Table 4.6 on page 79. pK_a values for all the titratable groups are shown, both for holo- and apo-BLA. The only large differences are found for the calcium coordinating carboxyl groups, D82, D87 and D88. There is also a relatively large difference for D84, which coordinates the calcium ion with its backbone oxygen.

4.6.6 New attempts at low pH

The effect of temperature and calcium concentration on the spectra was studied in order to try and obtain usable TOCSY spectra at lower pH than 3.9. The WATERGATE spectrum at pH 3.5 clearly shows that the chemical shift dispersion is limited (see Figure 4.18). Compared with spectra at higher pH, the linewidth are also larger, closer to 25 Hz relative to about 15 Hz at pH 6.0. The quality of the spectrum did not improved when calcium ions were added.

When lowering the temperature to 295 K, still at pH 3.5, the WATERGATE spectrum looked a lot better. The spectrum is not much different from the spectrum at 314 K and pH 3.9. At pH 3.2 the WATERGATE spectra at all the tested temperatures (283-314 K) had limited chemical shift dispersion.

A TOCSY spectrum was recorded at pH 3.5 and 295 K. Even though the WATERGATE spectrum at these conditions was good, the TOCSY spectrum was of poor quality. In the fingerprint region a lot of cross peaks could be found. However, most of them were weak and diffuse. Only a very limited number of them are possible to assign by comparing to the TOCSY spectra at higher pH values. In the region where the interesting H$^{\beta}$-H$^{\mathrm{N}}$ and H$^{\gamma}$-H$^{\mathrm{N}}$ cross peaks should be found, there are only a limited number of cross peaks. Moreover, none of the peaks that are possible to assign are relevant to the pK_a-determinations.

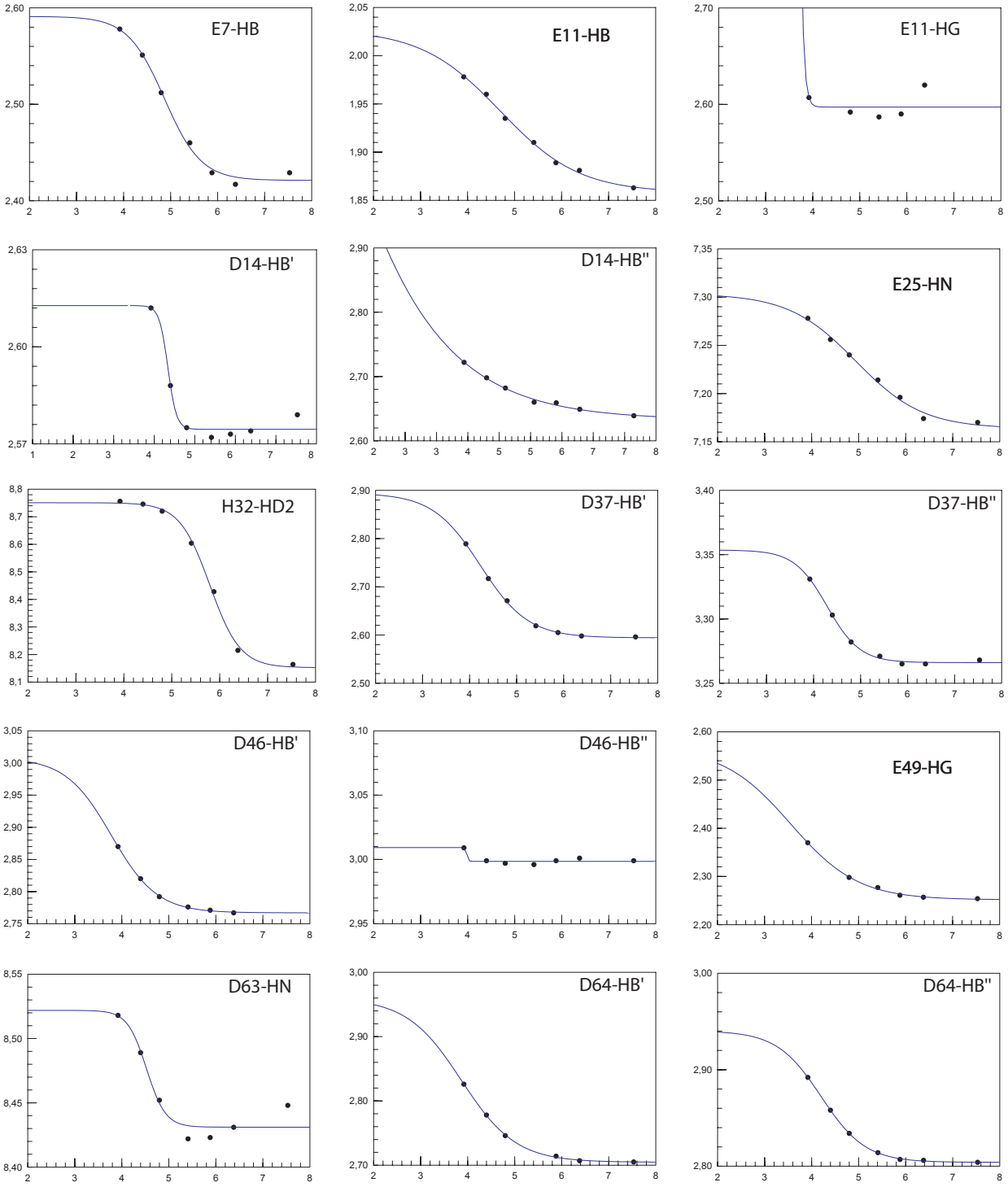


Figure 4.15: The figure shows the observed chemical shifts of the titrated residues from E7 to D64. The lines show the results of the regression analysis. The chemical shift is plotted on the y-axis, and the pH on the x-axis.

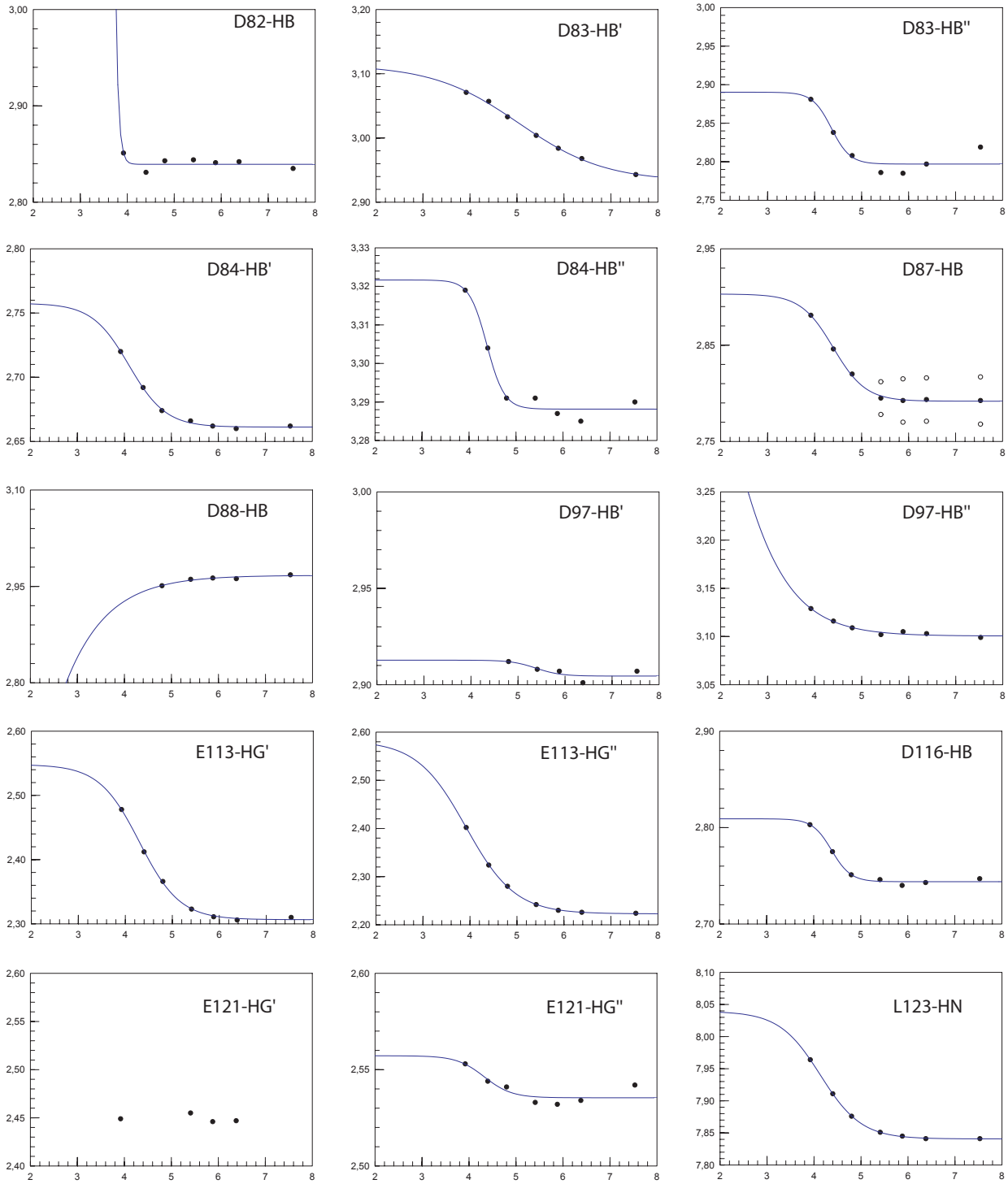


Figure 4.16: The figure shows the observed chemical shifts of the titrated residues from D78 to L123. The lines show the results of the regression analysis. The chemical shift is plotted on the y-axis, and the pH on the x-axis. For D87HB the open circles are the measured chemical shift, and the filled circles are the average values, which were used in the regression analysis. The regression analysis failed for E121HG'.

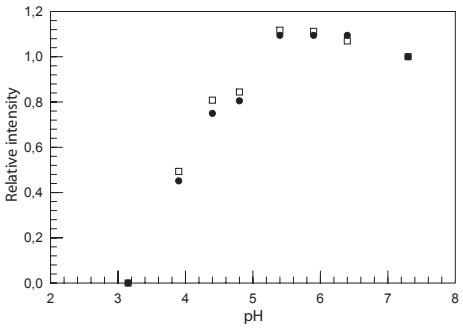


Figure 4.17: Intensities of V92H$^{\gamma 1/\gamma 2}$ and I95H$^{\delta 1/\delta 2}$D as a function of pH. The intensities are relative to the intensities of the signals from citric acid. The intensities at pH 7.53 are defined as 1. The WATERGATE spectra were used to measure the intensities.

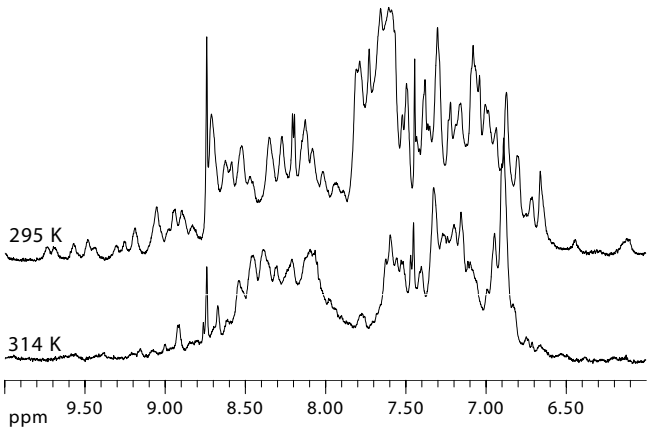


Figure 4.18: WATERGATE of BLA at pH 3.5 at 295 K and 314 K. The chemical shift dispersion is clearly considerably less at 314 K.

Table 4.6: Theoretical pK_a values

	holo-BLA		apo-BLA	
Resid.	pK_a	*n*	pK_a	*n*
N-term.	8.86	0.98	9.27	0.98
E1	4.51	0.74	4.84	0.63
K5	11.09	0.87	11.16	0.87
E7	4.56	0.78	4.40	0.76
R10	12.69	0.98	13.02	0.98
E11	4.68	0.74	4.73	0.67
K13	10.76	0.94	10.74	0.78
D14	3.67	0.90	3.66	0.90
K16	10.78	0.93	10.62	0.78
Y18	10.75	0.79	10.79	0.79
E25	3.74	0.90	3.66	0.90
H32	6.28	0.95	6.34	0.94
Y36	10.47	0.85	10.56	0.85
D37	3.56	0.85	3.56	0.82
D46	3.61	0.95	3.73	0.94
E49	4.48	0.90	4.52	0.88
Y50	11.71	0.91	12.04	0.92
K58	11.25	0.92	11.22	0.92
K62	10.81	0.94	10.85	0.94
D63	3.97	0.78	3.92	0.77
D64	4.07	0.89	4.03	0.90
H68	6.31	0.90	6.28	0.90
D78	3.86	0.85	3.99	0.83
K79	11.04	0.94	11.33	0.94
D82	2.47	0.72	4.44	0.61
D83	4.02	0.76	4.18	0.64
D84	3.73	0.79	4.31	0.70
D87	1.85	0.66	4.48	0.53
D88	1.98	0.65	4.47	0.54
K93	11.31	0.81	11.44	0.82
K94	10.86	0.82	10.98	0.83
D97	3.60	0.95	3.58	0.94
K98	10.77	0.88	10.79	0.88
Y103	10.92	0.85	10.52	0.88
H107	6.86	0.96	6.84	0.95
K108	10.66	0.97	10.71	0.96
E113	4.39	0.96	4.35	0.95
K114	10.69	0.96	10.72	0.97
D116	3.91	0.95	3.85	0.94
E121	4.51	0.96	4.45	0.96
K122	10.67	0.98	10.66	0.97
C-term.	3.03	0.92	3.03	0.91

Chapter 5

Discussion

5.1 ^{1}H-^{15}N HSQC versus FHSQC at natural abundance

According to Mori et al. [57] the advantage of using FHSQC compared with HSQC is that at short relaxation delay the signal intensities will be a lot better with FHSQC. This is in agreement with the results of the comparison done here (see section 4.1, page 59). The relaxation delay that was used is considerably less than the T_1 of water. Without having done a thorough inspection the most attenuated cross peaks in the HSQC spectrum seems to originate from solvent accessible residues. This happens because saturation is transferred from dephased water during the relaxation delay.

The signal-to-noise ratio of the FHSQC spectrum have been estimated to be about twice as good as that of the HSQC spectrum. In order to get spectra of equal quality the HSQC experiment would need four times as many scans as FHSQC. Alternatively a longer relaxation delay of the HSQC experiment would lead to a better spectrum.

If no other factors influence the choice of pulse sequence, the FHSQC experiment should be preferred instead of HSQC.

5.2 Assignment work

The assignment done here seems to be in good agreement with the assignment done by Forge et al. [31]. Five of the fingerprint resonances deviated by more than 0.5 ppm. These five resonances were T38H$^{\mathrm{N}}$, I41H$^{\alpha}$, S76H$^{\mathrm{N}}$, H107H$^{\alpha}$ and E121H$^{\alpha}$. The assignment of T38H$^{\mathrm{N}}$, I41H$^{\alpha}$ and S76H$^{\mathrm{N}}$ has to be regarded as correct because of the sequential NOEs confirming them. There are only one NOE correlation supporting the assignment of H107, and no NOEs for E121. On this basis these two assignments have to be regarded as very uncertain.

A few of the resonances that have been assigned by Forge et al. were not possible to assign here. This especially applied the N-terminal. None of the three first residues could be assigned. On the other hand, eight α-protons not assigned by Forge have been assigned here (see Table 4.1 on page 61).

5.2.1 Assignment of the H3 α-helix

The H3 α-helix was relatively easy to assign, including most of the side chains. There were also many NOE correlations, both sequential, medium-range (confirming that this

is in fact an α-helix) and long-range. This may indicate that the structure of the α-helix is very well defined. Alexandrescu et al. [27] have assigned several of the side chains in this helix. However, many of the assignments do not agree with those encountered here. The assignment by Alexandrescu et al. do not agree with Forge et al. [31] either. On this the assignment by Alexandrescu et al. is disregarded.

5.2.2 Assignment of the β-sheet

The sequential assignment of the β-sheet is based almost exclusively on $d_{\alpha\mathrm{N}}(i, i+1)$. This shows that the chain is extended and probably is a β-sheet. There is also a network of NOE correlations showing that it is an antiparallel β-sheet. There are very few correlations connecting the second β-strand (residue 47-50) to the third (residue 55-56). There are also only a couple of NOEs from the first and second β-strand to other parts of the protein. This can indicate that the β-sheet is very flexible.

5.2.3 Overlap problems

Overlapping is a severe problem in parts of the spectra. There are several regions of the fingerprint region where the cross peaks are impossible to separate. By combining the TOCSY, COSY and NOESY spectra most of them were fairly easy to assign. The biggest problems arose where peaks were impossible to separate in the COSY spectra. The cross peaks of L15H$^{\alpha}$-H$^{\mathrm{N}}$ and Y18H$^{\alpha}$-H$^{\mathrm{N}}$ were especially troublesome. In the COSY spectrum this looks like one peak, and only the use of ACME makes it possible to separate the two peaks. The fact that these two residues are close in space made the assignment even more difficult.

Overlapping in the H$^{\alpha}$-H$^{\beta}$ region made it impossible to assign many of the β-protons.

5.2.4 Abnormal chemical shift of I95H$^{\gamma 1}$

The very low chemical shift of one of the H$^{\gamma 1}$-protons of I95 is probably caused mainly by ring current effects from the six-membered aromatic ring of W60. This effect arises because the π-electrons of the aromatic ring set up a current to induce a magnetic field to counteract the changes of the magnetic flux through the aromatic ring. Protons on the outside of the ring will consequently experience a stronger magnetic field, while those on the "inside" will experience a weaker magnetic field.

In the x-ray structure of BLA (1F6S in the Protein Data Bank [4]) I95H$^{\gamma 12}$ is placed very close to the aromatic ring of W60. The distances to the aromatic protons are 3.7 Å, 3.8 Å, 3.7 Å and 3.5 Å to H$^{\epsilon 3}$, H$^{\zeta 3}$, H$^{\eta 2}$ and H$^{\zeta 2}$, respectively (see Figure 5.1). The expected ring current effect of W60 on I95H$^{\gamma 12}$ at this configuration is -2.73 ppm (calculated using the SHIFTCALC package [83]). I95H$^{\gamma 12}$ also has minor ring current effects from other nearby aromatic residues. The ring current effects from all nearby aromatic residues are -3.04 ppm. The expected chemical shift when including all contributions is -1.90 ppm. This is not too far from the measured chemical shift of -2.41 ppm. The explanation of the difference might be that there are contributions that SHIFTCALC does not include, or that the protein structure at the experimental conditions used here is slightly different from the x-ray structure.

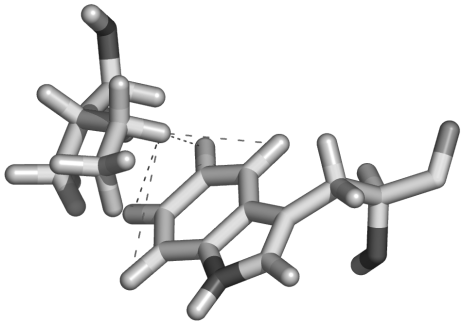


Figure 5.1: The position of I95 relative to W60. I95H$^{\gamma 2}$ is located almost directly above the aromatic ring. The ring current effects from the aromatic ring leads to an abnormally low chemical shift.

5.3 ACME

The software package ACME was used to measure ${}^{3}J_{\mathrm{H}^{\alpha}\mathrm{H}^{\mathrm{N}}}$ the coupling constants from the phase sensitive COSY spectrum. The measurement consisted of two steps. First a diagonal-only spectrum was used to estimate the intrinsic intensity, then the coupling constants were measured by restraining the intensity of the multiplets.

5.3.1 Estimation of intensity

Because of overlap on the diagonal only a few intensities were possible to measure. Ideally the intensity should be measured for each resonance. There is a large variation in the measured intensities. The reason why Y103H$^{\delta}$ has such a high intensity is that this signal is created by two protons, Y103H$^{\delta 1}$ and Y103H$^{\delta 2}$. Presaturation had to be used as water suppression in this experiment. The drawback of using presaturation is that exchangeable protons may be partially saturated due to saturation transfer. The resonances of these protons will have a low intensity.

The backbone amide protons of L85 and L23 have low intensities. According to the x-ray structure by Chrysina et al. [4] they are located on the surface of the protein, exposed to the solvent, and are therefore exchangeable. This also applies to the side chain amine proton of W60.

Forge et al. [31] have determined the protection factor of 49 of the backbone amide protons. This factor is a measurement of how fast the proton exchange is. Slow exchange equals large protection factor. They found both I75H$^{\mathrm{N}}$ and L96H$^{\mathrm{N}}$ to have a large protection factor. This explains their high intensity. C73H$^{\mathrm{N}}$ and G51H$^{\mathrm{N}}$ which have low intensities also have relatively low protection factors. The α-proton of N56 has a intensity similar to that of the non-exchangeable amide protons.

Another possible reason why the intensities may be low is saturation. If the recycling delay is too short the resonances with long T_1 will be saturated. However, the recycling delay used here is so long that this problem should be minimized.

It is unknown why D14H$^{\mathrm{N}}$ has such a high intensity. According to the x-ray structure this residue appears to be solvent accessible, and would be expected to have a lower intensity.

As a compromise, the value of 1000 was chosen as the intrinsic intensity used in the measurement of the coupling constants

5.3.2 Measurement of coupling constants

Many of the cross peaks found in the DQF-COSY spectrum are missing in the COSY spectrum used with ACME, especially cross peaks close to the water resonance. This problem is due to the solvent suppression method used. The presaturation not only saturates the water resonance, but it also affects resonances close to this frequency, the α-protons. Another reason why cross peaks are not found in the COSY spectra is their coupling constant. Cross peaks with small coupling constants will be considerably weaker than those with large coupling constant.

The study of the effect of different intrinsic intensities showed that a too low intensity leads to a too large coupling constant (see Figure 4.9 on page 68). In the measurement of the coupling constants the same intrinsic intensity has been used for all resonances. The consequence of this is that less protected residues, residues with a low intrinsic intensity, are given a too low coupling constant. Highly protected residues are given a too high coupling constant. The concrete effect is that the measured coupling constants of the residues in the β-sheet, which is solvent accessible, are smaller than the real coupling constants. The coupling constants of the residues in the highly protected H3 α-helix are in reality smaller than measured.

A good idea would be to use different intrinsic intensities for the different parts of the proteins. For the highly protected H3 α-helix an intrinsic intensity of about 1100 could be used, while 700-800 could be used as the intrinsic intensity for clearly solvent accessible residues. A more time consuming alternative could obviously be to measure the protection factor of all amide protons. Anyway, the partial saturation of certain α-protons would still be a complicating factor.

Figure 5.2 shows the measured coupling constants as a function of the ϕ-angle. The angles are taken from the x-ray structure by Chrysina et al. [4]. The figure also shows the Karplus curve determined by Pardi et al. [41]. This was the curve used to derive the dihedral angles in this thesis. Taking into account the estimated uncertainty of 1 Hz, there is a satisfactory agreement between the measured coupling constants and the dihedral angles from the x-ray structure. The largest deviations originate from residues with dihedral angle less than -100°. These residues are located in the β-sheet and C-terminal. They residues are solvent accessible, and therefore probably have an intrinsic intensity considerably less than the estimated intrinsic intensity. Motional averaging could be another explanation. It should be noted that the x-ray structure used to assess the quality of the measured coupling constants probably is not the same as the structure in solution under the experimental conditions used here.

If coupling constants are not of interest, ACME should still be used as a tool to separate overlapping cross peaks in the COSY spectrum.

5.4 CSI

The chemical shift index provides a simple yet powerful tool to determine the secondary structures of a protein, and hence the two backbone torsion angles.

If we compare the results of CSI with the secondary structure of the x-ray structure we find that there is a good correlation (see Figure 1.10 on page 19, and Figure 4.10 on page 68). CSI correctly determines that there is a large α-helical region in the N-terminal. The three β-strands are also detected by CSI, even though the length of the third strand is slightly exaggerated. CSI also predicts two β-strands in the C-terminal. These are not found in the x-ray structure, and no other experimental data confirms

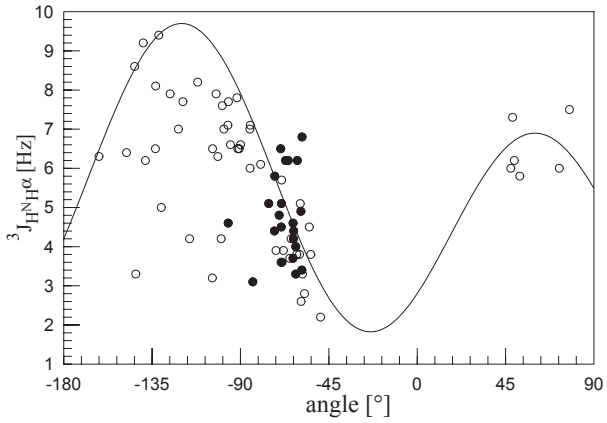


Figure 5.2: Plot of the coupling constants measured with ACME as a function of dihedral angle. Residues in the α-helices are marked by closed circles. The dihedral angles are taken from the x-ray structure of BLA determined by Chrysina et al. [4]. The Karplus curve determined by Pardi et al. [41] is also shown.

them. Although this could be a structural feature just present in the protein in solution, it is rather believed that they are a reflection of the imperfections of CSI.

According to CSI the residues 68-74 are in an α-helical conformation. According to the x-ray structure this is an unstructured region. The coupling constants of N71, C73 and N74 are also too large for this to be an α-helix. CSI correctly predicts the H3 α-helix, and the two small 3_{10}-helices in the C-terminal. In the x-ray structure residues 105-110 are also an α-helix. There are several reasons to why this is not an α-helix according to CSI. First of all too many of the α-proton chemical shifts are missing. The assigned H^α-resonance of H107 could also be wrong, as also discussed in section 5.2. To indicate an α-helix chemical shift should be less than 4.73 ppm, while it is assigned to be 4.83 ppm. The region composed of residues 105-110 has been reported to both being able to adopt a helical and a loop structure [84] in human α-LA. At low pH this region adopts a highly flexible loop structure. The pH used here, pH 5.6, is probably low enough for this region to be a flexible loop. The reason why CSI does not predict an α-helix for the residues 105-110 is most likely a combination of the three explanations mentioned.

Overall 73% of structures determined by CSI agree with the x-ray structure. This is low compared with the expected 90-95% accuracy if the structures were similar in the crystal and in solution. Possible reasons for the lower accuracy include missing chemical shifts, wrong assignments, different experimental conditions, and that the protein is very flexible.

5.5 Structural calculations

In the structural calculations two equal NOESY spectra were used, one manually peak picked as a part the assignment work, and one automatically peak picked using AUTOPSY. A consequence of this is that there are duplicates of many of the restraints. The manually peak picked spectrum contained mostly intraresidual and sequential NOEs, and a few medium-range. Compared with Table 4.4 on page 69 the number of unique

intraresidual, sequential and medium-range restraints are fewer, probably closer to half of the values in the table. Most of the long-range restraints are assumed to be unique.

5.5.1 Large RMSDs

Several regions of the calculated structure have a very large RMSDs, most notably the terminals, and regions close to residue 20, 45 and 65. There are several explanations for these large RMSDs, and they all have to do with the lack of NOE correlations. The assignment of the proton resonances of BLA achieved in this work is not complete. As a consequence several important NOE correlations might have been overlooked. Overlapping of NOESY cross peaks also creates a problem. This problem is partly overcome by the use of ADRs, but in parts of the spectrum the overlapping is too severe even for ADRs. The most problematic regions are the H^{α}-H^{β}-region and the regions close to the diagonal, especially between 0.5 ppm and 2.5 ppm. Conformational flexibility of the protein is another probable reason to why there are few NOE correlations for some parts of the protein.

5.5.2 Fluctuating structure

The structural determinations of BLA by x-ray crystallography [4, 85] have shown that there is considerable flexibility in parts of the protein. In x-ray crystallography the temperature factor, the B-factor, is used as a measurement of mobility. The B-factor can be related to the mean displacement of a vibrating atom. In addition to high B-factors for the terminals, especially the C-terminal, high B-factors are also reported for the residues 43-47 and 62-65. Pike et al. [85] also report elevated B-factors for the residues around G20. At low pH Harata et al. [84] have also reported an increased mobility of the region L105-L110 in human α-LA. These are exactly the regions with high RMSD values, which indicate that flexibility and true mobility are partly responsible for the high RMSDs.

Also, the region consisting of L105 to L110 is missing a lot of resonances. The reason for this is, according to Forge et al. [31], that the cross peaks are broadened beyond detection, possibly as a result of flexibility.

5.5.3 The quality of the calculated structure

The ring current effects on I95HH$^{\gamma 12}$ have been discussed briefly in section 5.2.4. In the x-ray structure, PDB 1F6S, the ring current effects from W60 is -2.73 ppm, and the ring current effects from other sources is -0.31 ppm. In the calculated structure I95H$^{\gamma 12}$ is placed a little more away from, and to the side of, the aromatic ring of W60, leading to a smaller ring current effect, -2.43 ppm. However, the biggest difference originates from other ring current sources, +0.30 ppm, relative to -0.31 ppm in the x-ray structure. The estimated chemical shift of I95H$^{\gamma 12}$ in the NMR structure obtained in this work is -0.99 ppm. This is far from the observed chemical shift, -2.41 ppm. The large difference between the observed and the calculated chemical shift is a strong indication that the calculated structure is not totally correct at this region.

The most obvious difference between the x-ray structure and the calculated structure is that the cleft region is missing. There were almost no NOE restraints that could restrain the position of the flexible loop (residues L105-L110) and the β-sheet (residues I41-Y50). A consequence of this is that these two parts of the protein are so far apart that there is no cleft. Because the cleft is believed to be important for the function of

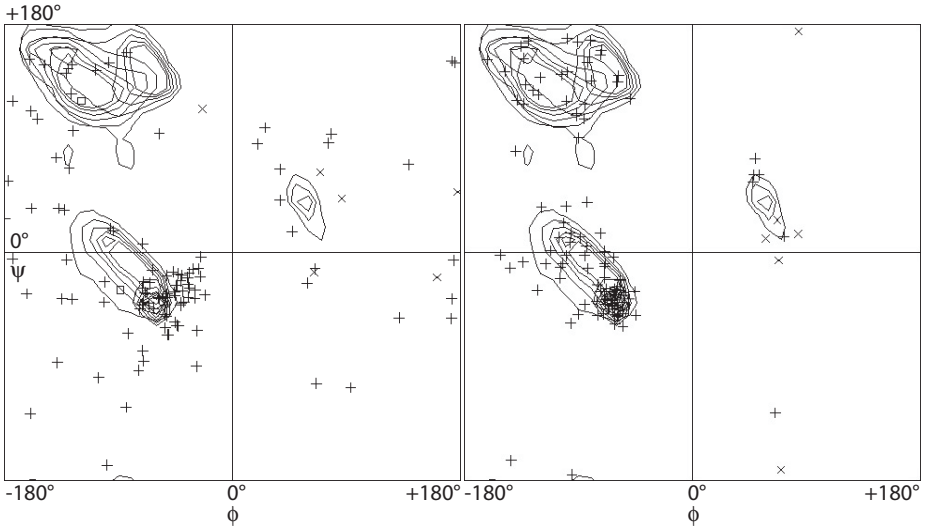


Figure 5.3: Ramachandran plot of the calculated structure (left) and the x-ray structure [4] (right). The crosses show the ϕ and ψ angles of the residues. The contour show the ϕ and ψ angles which give the lowest potential energy. It is evident that the x-ray structure has a more energetically favorable conformation. The large accumulation of residues in the lower left quadrant are residues in the α-helices. These are well structured in the calculated structure.

lactose synthase [86], the flexible loop and the β-sheet is probably closer together in the true structure.

The Ramachandran plot of the calculated structure also show that the quality is not too good (see Figure 5.3). This plot shows which ϕ and ψ angles give the lowest potential energy, and what the ϕ and ψ angles of the structure are.

5.5.4 Structure and function

A mutagenesis study of BLA has shown that L110, and partly A106, both on the flexible loop (residue L105-L110), are important for the functions of BLA [86]. Together with F31, H32, Q117 and W118 they form a surface that is crucial for the function of lactose synthase. Another study has shown that a mutation of K114 which lead to increased stability of this region, also decreased the affinity of BLA for GT [87]. This is seen as a proof that the flexibility of this region of the C-terminal is important to the function of BLA.

An x-ray structure of the lactose synthase complex shows that the aforementioned residues are in close contact with GT [88]. In addition to these, E2, V42 and D44 are also in close contact. These three residues are also located on parts of the chain which have large flexibility according to the B-factor and the RMSDs found in this study. This may suggest that the flexibility of these residues also are important for the function.

5.5.5 Changes that could be done to the data

The calculated structure violates many of the dihedral restraints derived from CSI and the ${}^3J_{\mathrm{H}^\alpha\mathrm{H}^\mathrm{N}}$-coupling constants. One reason for this is that the information from CSI has been trusted too much. It would probably have been a good idea to remove more

of these restraints. The restraints that should be removed are the following α-helices: L52-I55, H68-N74 and D116-C120. Almost no NOE correlations support the fact that these actually are α-helices.

As for the coupling constants, all the sources of error mentioned in section 5.3 probably makes the uncertainty greater than the estimated 1 Hz. A solution could be to only use the coupling constant as an indication of α-helix or β-sheet.

The possibility to disregard these dihedral restrains in the structural calculations should not be considered as an option. This was briefly tried, and the result was a poorly defined structure.

It is also possible to experiment with different parameters for the structural calculations. This was not tested extensively here, as the structural calculations are very time consuming.

5.5.6 What could be expected of structures from NMR data?

How good structures can we expect from NMR? A low detail structure have RMSDs of 1.5-2.0 Å for the backbone atoms. A structure with RMSD less than 1.0 Å is regarded as a good structure. The structure determined here has an average RMSD of 1.6 Å for all backbone atoms, 1.1 Å for backbone atoms excluding the terminals, and 0.4 Å for the H3 α-helix.

Schwalbe et al. [89] have determined a very good structure of hen lysozyme, a protein homologous with BLA, with an average RMSD of 0.5 Å for the backbone atoms. This structure was obtained using a doubly labelled protein (^{13}C and ^{15}N), and residual dipolar couplings (RDCs). An older study of the same protein, using only homonuclear methods, achieved an RMSD of 1.7 Å for the backbone atoms [90].

Hen lysozyme is a relatively stable protein relative to the fluctuating BLA. An x-ray structure of hen lysozyme reported an average temperature factor of 10 Å^2, with peaks close to 30 Å^2 [91]. The average temperature factors for BLA is 45 Å^2, with peaks close to 80 Å^2.

5.6 pK_a determinations

5.6.1 pK_a determinations using BASHD-TOCSY

The BASHD-TOCSY spectra gave considerably worse spectra than ordinary TOCSY. Most cross peaks which originated from protons separated by 3 bonds gave good cross peaks in the BASHD-TOCSY spectra. Protons separated by 4 or more bonds generally gave no, or poor, cross peaks. The exception seemed to be cross peaks from protons which both had been effected by the selective pulse. Because the pK_a determinations relied on the H$^\beta$-H$^\mathrm{N}$ and H$^\gamma$-H$^\mathrm{N}$ cross peaks this made BASHD-TOCSY useless.

The experimentation with different acquisition parameters did not give better spectra. But it is unknown whether this is a general problem with BASHD-TOCSY, or if it is the acquisition parameters used here which create the problem.

5.6.2 pK_a determinations using ordinary TOCSY

The ordinary TOCSY experiment gave decent spectra as low as pH 3.92. In addition to being a tool for the pK_a determinations, the TOCSY spectra at different pH also aided in the assignment. Several ambiguous cross peaks were assigned because of small pH

variations in the chemical shifts. The experimentally determined pK_a values will be compared with theoretical pK_a values (see also Table 6.1 on page 96).

Some of the residues have a Hill coefficient higher than 1, this may reflect a pK_a value which is too high. D87 has a Hill coefficient of 1.3. If this was restrained to 1.0 the pK_a value would be 0.1 units lower.

5.6.3 Residues with good titration curves

E7 As the pH changed it was revealed that the cross peaks assigned as E7H$^{\beta}$ was actually two cross peaks. It is unknown whether these cross peaks were H$^{\beta 2}$ and H$^{\beta 3}$, or possibly H$^{\beta 2/\beta 3}$ and H$^{\gamma 2/\gamma 3}$. Only one of the cross peaks had a sigmoidal behavior. This was the one used to determine the pK_a value. pK_a for this residue was determined to be 4.9$\pm$0.1, slightly higher than the theoretical value, 4.56. A possible explanation for the difference can be the flexibility of this part of the protein.

E11 The γ-protons did not give a titration curve, but the regression analysis of the β-protons gave a sensible result. The experimentally determined pK_a value, 4.7$\pm$0.3, is exactly the same as theoretical value. Both the experimental and theoretical pK_a determination give this residue a low Hill coefficient. This can probably be attributed to the close contact with E7.

E25 The chemical shift of the backbone amide proton was used to determine the pK_a, as the chemical shifts of the β- and γ-protons were unknown. The regression analysis gave a high pK_a value, 4.9$\pm$0.3, and a low Hill coefficient, 0.6$\pm$0.2. The proximity to H107 is probably the reason for the low Hill coefficient. The theoretical calculation gave a much lower pK_a, 3.74. The reason for this large difference is that the theoretical calculation has not taken into account the structural changes of the flexible loop (residue L105-L110). The theoretical calculations have assumed an ionic bond between H107 and E25, but this ionic bond is broken by the structural changes around this region due to the assumed flexibility.

D37 Both β-protons of this residue gave reliable pK_a values. 4.2$\pm$0.1 and 4.3$\pm$0.1 with the Hill coefficients 0.86$\pm$0.08 and 1.2$\pm$0.3 respectively. Based on the uncertainties and correlation coefficients 4.2$\pm$0.1 is believed to be the more correct pK_a value. The theoretical calculation gave a lower pK_a value, 3.56, with the Hill coefficient 0.85. The difference may be explained by the flexibility of the nearby N-terminal.

D46 Only one of the β-protons of D46 exhibits a sigmoidal behavior, the other chemical shift is virtually independent on pH. The determined pK_a value, 3.8$\pm$0.3, is in accordance with the theoretical value, 3.61. The high correlation coefficient and the low uncertainties also shows that it is a trustable value. The reason for why the cross peaks have disappeared at pH 7.53 is unknown.

E49 The uncertainty of the pK_a value of E49 is very large: 4$\pm$1. Due to this reason this value is not supposed to be very faithful. The theoretical value is 4.48.

D64 Both the β-protons of D64 give nice looking titration curves. The correlation coefficient is high and the uncertainties are low, but the two pK_a values are a little different, 3.9$\pm$0.3 and 4.20$\pm$0.06. A sensible estimation of the true pK_a value is 4.1$\pm$0.2. This is the same as the theoretical value, 4.07.

D83 The β-protons give very different results. The one called HB' have a quite high pK_a value, 5.0$\pm$0.2, and a low Hill coefficient, 0.47$\pm$0.08. Based on the x-ray structure, the low Hill coefficient can partly be explained by the closeness to E1. The other proton have a lower pK_a and a higher Hill coefficient, 4.4$\pm$0.3 and 2$\pm$3. There seems to be sources other than protonation that affect the chemical shift of this proton. A conclusion about the true pK_a is hard to draw, but it could be estimated to 4.5$\pm$0.5. The theoretically determined pK_a value is 4.02.

D84 The two pK_a values determined for the β-protons of D84 are not in accordance with each other either. The proton with the highest correlation coefficient, pK_a=4.1$\pm$0.2, is assumed to be the most correct one. This value is slightly higher than the theoretically calculated pK_a value, 3.73.

E113 Both the titration curves of E113 give good correlation coefficient and low uncertainties, but they do not agree with each other. A compromise between the two would be 4.1$\pm$0.3. The theoretical value is 4.39.

L123 The C-terminal carboxyl group was the group that gave the best fit. The experimentally determined pK_a value, 4.14$\pm$0.05, is higher than the theoretical value, 3.03.

5.6.4 Residues with bad titration curves

D14 The uncertainties of the coefficients from the regression analysis are very large for D14. A likely explanation is that at the lowest pH value the protonation of this residue have just started. This means that the pK_a value is less than 3.9, a reasonable estimation is 3.5$\pm$0.3. This is in agreement with the theoretical value, 3.67.

D63 At high pH there seems to be other sources which affect the chemical shift. The use of HN, in stead of the β-protons, makes the results uncertain. The theoretical pK_a value, 3.97, is quite far from the experimentally determined, 4.5$\pm$0.2.

D116 The Hill coefficient of D116 is higher than it should be, 2.1$\pm$0.7. Based on the fact that the expected chemical shift difference between the fully protonated and fully deprotonated should be around 0.2 ppm, it appears that the protonation has just started. As with D14 the expected pK_a value is estimated to be 3.5$\pm$0.3, which is lower than the theoretical value, 3.91.

5.6.5 Residues without a titration curve

D97 As with D14 and D116 the protonation of D97 seems to just have started at pH 3.9. In agreement with what has been estimated for D14 and D116 the pK_a is estimated to be 3.5$\pm$0.3. The theoretically calculated pK_a, 3.60, is within this range.

E121 There are no observed titration curve for the two γ-protons of E121. There are two possibilities: either the pK_a value is too low to provide any chemical shift changes in the pH range studied, or the assignment of this residue is wrong. Based on the fact that this residue has no NOE correlations with other residues (it was assigned on a closest match basis), and the fact that the theoretical pK_a for this residue is 4.51, it is most likely that the assignment is wrong. The regression analysis of HG' failed because there were only four points to which the equation should be fitted to. Because the equation has four unknowns, more than four points are needed.

5.6.6 The calcium binding carboxyl groups residues

Of the carboxyl groups which coordinate the calcium ion, D82, D87 and D88, only D87 exhibit a titration curve. At low pH the two β-protons of D87 have the same chemical shift, but at high pH they are slightly different. The average of the two were used to calculate the pK_a. The determined pK_a is quite high to be an aspartate residue, 4.40$\pm$0.08. As a comparison the theoretical value, in presence of the calcium ion, is 1.85. The other two residues, D82 and D88, do not appear to become protonated, probably because of the calcium ion.

The protonation of only D87 is surprising considering that D82 is in an almost equal chemical environment as D87 according to the x-ray structure of holo-BLA. This is also reflected by the higher theoretical pK_a for D82 than D87. A question could then be asked: Does this really reflect the calcium bound state? The following facts are known:

- The temperature used is probably higher than the melting temperature, T_m, for the apo-protein under the conditions used here (T_m=33.9 °C in presence of 0.05 M NaCl [92]). This would lead to a poorly resolved NMR spectrum, with possibly undetectable cross peaks.
- The intensity of the resonances are drastically decreasing (see Figure 4.17 on page 78), possibly reflecting the decreasing fraction of holo-BLA, and increasing fraction of apo-BLA
- The rate constant for conversion of holo to apo is low compared with the NMR time scale (0.37$\pm$0.04 s^{-1} at 20 °C [32]). This will lead to two sets of signals, one for apo and one for holo.
- The chemical shifts of the calcium binding residues is completely different in apo than in holo [32]. The chemical shifts observed here are closest to those reported for holo-BLA.

Based on this the most likely explanation is that observations only reflect the calcium bounded state.

It is still unknown why D87 becomes protonated and not D82. A possible explanation may be that there are small structural changes in the side chains as the pH decreases. When D87 is protonated the water molecule between D82 and D87 is probably lost (see Figure 5.4). As a compensation for the loss of two hydrogen bonds, an ionic bridge might be formed between the carboxyl group of D82 and the amino group of K79. In the case of protonation of D82 the formation of an ionic bridge between D87 and K79 is unlikely because of sterical hindrance, which leads to an energetically unfavorable situation. The protonation of both residues will most likely lead to loss of the calcium ion.

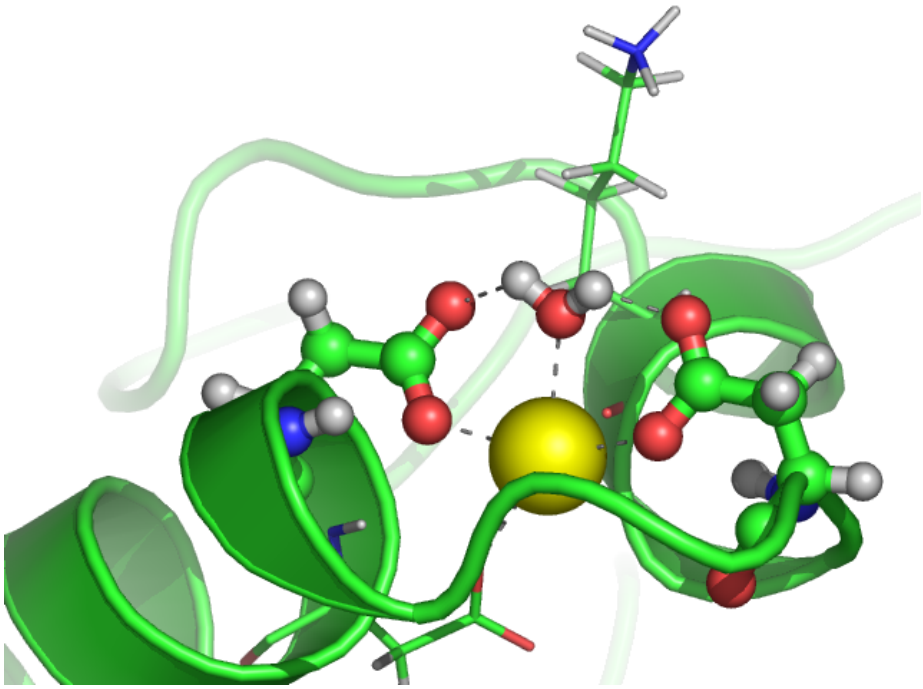


Figure 5.4: The calcium binding site of BLA. The protonation of D87 (left) probably leads to loss of the coordinating water molecule. The protonation of D87 and not D82 (right) might result in the formation of an ionic bridge between K79 (top) and D82.

The lesser importance of the negatively charge of D87 for maintaining the structure of the holoprotein is also confirmed by a mutagenesis study. D87 was replaced by an asparagine residue, without significant decrease in temperature stability (~57 °C versus ~66 °C) [14].

5.6.7 The histidine residues

The pK_a values of the histidine residues have earlier been determined by Alexandrescu et al. [30]. The values they obtained was 5.28/5.18 for H32, 6.23/6.30 for H68 and 6.28/6.24 for H107.

H32 The pK_a value determined for H32 is quite low, only 5.79$\pm$0.04. The theoretical pK_a value is 6.28. This large difference is most likely due to the structural changes for the flexible loop (L105-L110). The low pK_a value could indicate that H32 becomes more buried among hydrophobic residues, making protonation unfavorable. H32 is important with respect to the function of lactose synthase, but it is unknown whether the low pK_a value is of any importance. This pK_a value may not only reflect the protonation, but also the structural changes.

H107 The structural changes of the region in which H107 is located makes the theoretical pK_a different from the true pK_a. The experimentally determined pK_a value for H107 is 6.02$\pm$0.05. This is a much lower than the theoretically calculated pK_a value,

6.86. In the x-ray structure used to calculate the theoretical pK_a there is the possibility of forming an ionic bond with E25, which makes protonation favorable. However, structural changes making H107 more solvent accessible destroy this possibility.

5.6.8 pK_a and membrane interaction

Holo-BLA interacts with negatively charged membranes through electrostatic forces [1]. The mechanism for membrane interaction proposed by Halskau et al. [2] is dependent on the protonation of D97, and also partly D14. Considering that membrane interaction occurs at pH 4.5, and that pK_a of D97 and D14 appear to be considerably lower, around 3.5, this does not at first seem likely. But considering that the pH close to a negatively charged membrane is lower than in the bulk solution, it is still a possibility.

It is interesting to notice that many of the carboxyl groups with pK_a close to 4.5 are located in the same region, around the N-terminal (E1, E7, E11 and D37) and the calcium binding site (D83, D84 and D87). At pH 5.5 this region is highly negatively charged, while it is estimated to be slightly positively charged at pH 4.5. At lower pH it gets even more positive. Other regions of the protein have only minor changes in charge. This drastic change of charge for only this region will considerably change the dipole moment of the protein. This change is possibly what is needed to correctly orient the protein over membrane. After this initial change in orientation electrostatic interaction might bring the protein close enough to the membrane for hydrophobic interaction between the protein and the membrane to occur.

5.6.9 New attempts at low pH

The effect of temperature and calcium concentration on the spectra was studied in order to try to obtain usable TOCSY spectra at lower pH than 3.9. The limited chemical shift dispersion observed at 314 K and pH 3.5 suggest the appearance of the molten globule conformation. It might in fact the premolten globule proposed by Gussakovsky and Haas [26].

At 295 K it was possible to obtain spectra with a normal chemical shift dispersion, but the TOCSY spectrum recorded at this temperature was of poor quality. A reasonably explanation is that only apo-BLA is present under these conditions.

Based on this there seems to be no use in performing a new pH titration at lower temperature. The additional information gained by doing this would probably very limited.

5.7 Further work

- The dipole moment (magnitude and direction) of BLA should be calculated, at least at pH 4.5 and at pH 6.0, to determine whether this is important for membrane interaction.
- The effect of protonation of D82 and D87 should be studied by computational methods in order to determine why D87 is protonated and not D82.
- Do a pH titration of apo-BLA in order to determine the pK_a without calcium present. Large changes in the pK_a values of the calcium binding site would then be expected, especially for the carboxyl group which coordinates the calcium ion.

- Try and confirm the location of the strong zinc binding site of BLA. There are at least two methods this could be done using NMR. A pH titration could be done in the presence of zinc ions. If the binding site involve E1, E7, D11 and D37, as proposed by Permyakov et al. [10], changes in the pK_a values of these residues would be expected. Alternatively the chemical shifts of these residues could be observed in the presence and absence of zinc ions. Considerable changes of their chemical shifts would indicate that this was the strong zinc binding site.

- Isotopic labelling of the protein is necessary. This would simplify the assignment process considerably, and make it possible to get a more complete assignment. The use of three-dimensional experiments would also make the overlapping of NOE correlation an insignificant problem. Labelling would also lead to new methods of measuring the $^3J_{\mathrm{H}^\alpha\mathrm{H}^\mathrm{N}}$-coupling constants. The results from CSI, or equivalent software packages, would be more reliable when also including ^{13}C chemical shifts. Isotopic labelling would also give alternative chemical shifts to use in pK_a determinations.

- Solide state NMR should be used to determine the position and conformation of BLA in membrane.

Chapter 6

Conclusions

6.1 Structural calculations

The structural calculations of BLA give structures with relatively low level of detail. The helical regions have a relatively good precision, but there are four larger regions of the protein which have poorly defined structure. These regions are both the terminals in addition to the β-sheet and the region close to N66.

The main reason for these poorly defined regions is most likely the flexible nature of the protein. This makes it impossible to obtain a structure of the protein in solution with both high precision and high accuracy.

The high flexibility is probably the main reason to why there are no NMR structures of BLA, or any other α-LA, in the RCSB Protein Data Bank yet.

6.2 pK_a determinations

The flexibility and dynamics of the protein create difficulties for the pK_a determinations. Because of conformational changes around pH 3.7 the end point of the titration has not been reached for many of the residues. This is the reason for the high uncertainties.

The conformational changes also creates problems for the theoretical calculation of the pK_a values. The conformational changes are not taken into account, which makes the pK_a values wrong. Notably, the conformational changes of the flexible loop (L105-L110) seems to be responsible for the wrong theoretical determination of the pK_a values for H32, E25 and H107.

D97 which has been implicated in membrane interaction has a low pK_a value, only around 3.5. This may indicate that protonation of this residue is not important for the initial step of the membrane interaction.

Many of the residues with pK_a around 4.5 are located in the same region of space, close to the N-terminal (E1, E7, E11 and D37) and the calcium binding site (D83, D84 and D87). A consequence of this is that around this pH there will be a large change in the dipole moment of the protein. This change may be what is needed to correctly orient the protein over the membrane. After this change of orientation electrostatic interactions might bring the protein close enough to the membrane for hydrophobic interaction to occur between the protein and the membrane.

A summary of the determined pK_a values can be found in Table 6.1.

Table 6.1: Summary of the pK_a values determined. Some of the values are not available (*NA*).

	theoretical		experimental		
Resid.	pK_a	*n*	pK_a	*n*	note
NT	8.86	0.98	*NA*	*NA*	
E1	4.51	0.74	*NA*	*NA*	
E7	4.56	0.78	4.9±0.1	1.1±0.2	
E11	4.68	0.74	4.7±0.3	0.5±0.1	
D14	3.67	0.90	3.5±0.3	*NA*	Experimental: uncertain
E25	3.74	0.90	4.9±0.3	0.6±0.2	Theoretical: wrong, experimental: uncertain
H32	6.28	0.95	5.79±0.04	1.4±0.2	Theoretical: wrong
D37	3.56	0.85	4.2±0.1	0.86±0.08	
D46	3.61	0.95	3.8±0.3	0.9±0.1	
E49	4.48	0.90	4±1	0.6±0.2	
D63	3.97	0.78	4.5±0.2	2±1	Experimental: uncertain
D64	4.07	0.89	4.1±0.2	0.9±0.1	
H68	6.31	0.90	*NA*	*NA*	
D78	3.86	0.85	*NA*	*NA*	
D82	2.47	0.72	*NA*	*NA*	Not protonated in holo-BLA
D83	4.02	0.76	4.5±0.5	*NA*	
D84	3.73	0.79	4.1±0.2	1.1±0.2	
D87	1.85	0.66	4.40±0.08	1.3±0.2	
D88	1.98	0.65	*NA*	*NA*	Not protonated in holo-BLA
D97	3.60	0.95	3.5±0.3	*NA*	Experimental: uncertain
H107	6.86	0.96	6.02±0.05	1.0±0.1	Theoretical: wrong
E113	4.39	0.96	4.1±0.3	0.9±0.1	Experimental: uncertain
D116	3.91	0.95	3.5±0.3	*NA*	Experimental: uncertain
E121	4.51	0.96	*NA*	*NA*	Probably wrong assignment
L123	3.03	0.92	4.14±0.05	0.99±0.04	

Bibliography

[1] S. Bañuelos and A. Muga. Structural requirements for the association of native and partially folded conformations of α-lactalbumin with model membranes. *Biochemistry* **35**, 3892–3898 (1996).

[2] O. Halskau, N. A. Frøystein, A. Muga, and A. Martínez. The Membrane-bound Conformation of α-Lactalbumin Studied by NMR-monitored ^{1}H Exchange. *J. Mol. Biol.* **321**, 99–110 (2002).

[3] J. L. Markley, A. Bax, Y. Arata, C. W. Hilbers, R. Kaptein, B. D. Sykes, P. E. Wright, and K. Wuthrich. Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J. Mol. Biol.* **280**, 933–952 (1998).

[4] E. D. Chrysina, K. Brew, and K. R. Acharya. Crystal structures of apo- and holo-bovine α-lactalbumin at 2.2 Å resolution reveal an effect of calcium on inter-lobe interactions. *J. Biol. Chem.* **275**, 37021–37029 (2000).

[5] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

[6] W. L. Delano. The PyMOL Molecular Graphics System. *Delano Scientific* (2002).

[7] K. Brew, T. C. Vanaman, and R. L. Hill. Comparison of the Amino Acid Sequence of Bovine α-Lactalbumin and Hens Egg White Lysozyme. *J. Biol. Chem.* **242**, 3747–3748 (1967).

[8] K. R. Acharya, D. I. Stuart, N. P. C. Walker, M. Lewis, and D. C. Phillips. Refined Structure of Baboon α-lactalbumin at 1.7 Å Resolution - Comparison with C-Type Lysozyme. *J. Mol. Biol.* **208**, 99–127 (1989).

[9] G. Vriend. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.* **8**, 52–56 (1990).

[10] E. A. Permyakov and L. J. Berliner. α-Lactalbumin: structure and function. *FEBS Lett.* **473**, 269–274 (2000).

[11] D. B. Veprintsev, M. Narayan, S. E. Permyakov, V. N. Uversky, C. L. Brooks, A. M. Cherskaya, E. A. Permyakov, and L. J. Berliner. Fine tuning the N-terminus of a calcium binding protein: α-lactalbumin. *Proteins: Struct., Funct., Genet.* **37**, 65–72 (1999).

[12] E. A. Permyakov, V. V. Yarmolenko, L. P. Kalinichenko, L. A. Morozova, and E. A. Burstein. Calcium-Binding to α-Lactalbumin - Structural Rearrangement and Association Constant Evaluation by Means of Intrinsic Protein Fluorescence Changes. *Biochem. Biophys. Res. Commun.* **100**, 191–197 (1981).

[13] E. A. Permyakov, L. A. Morozova, and E. A. Burstein. Cation binding effects on the pH, thermal and urea denaturation transitions in α-lactalbumin. *Biophys. Chem.* **21**, 21–31 (1985).

[14] S. E. Permyakov, V. N. Uversky, D. B. Veprintsev, A. M. Cherskaya, C. L. Brooks, E. A. Permyakov, and L. J. Berliner. Mutating aspartate in the calcium-binding site of α-lactalbumin: effects on the protein stability and cation binding. *Protein Eng.* **14**, 785–789 (2001).

[15] N. Chandra, K. Brew, and K. R. Acharya. Structural evidence for the presence of a secondary calcium binding site in human α-lactalbumin. *Biochemistry* **37**, 4767–4772 (1998).

[16] J. S. Ren, D. I. Stuart, and K. R. Acharya. α-Lactalbumin Possesses a Distinct Zinc-Binding Site. *J. Biol. Chem.* **268**, 19292–19298 (1993).

[17] S. E. Permyakov, D. B. Veprintsev, C. L. Brooks, E. A. Permyakov, and L. J. Berliner. Zinc binding in bovine α-lactalbumin: Sequence homology may not be a predictor of subtle functional features. *Proteins-Structure Function and Genetics* **40**, 106–111 (2000).

[18] J. M. Aramini, T. Hiraoki, M. R. Grace, T. W. Swaddle, E. Chiancone, and H. J. Vogel. NMR and stopped-flow studies of metal ion binding to α-lactalbumins. *Biochim. Biophys. Acta* **1293**, 72–82 (1996).

[19] A. Pellegrini, U. Thomas, N. Bramaz, P. Hunziker, and R. von Fellenberg. Isolation and identification of three bactericidal domains in the bovine α-lactalbumin molecule. *Biochim. Biophys. Acta* **1426**, 439–448 (1999).

[20] A. Håkansson, M. Svensson, A. K. Mossberg, H. Sabharwal, S. Linse, I. Lazou, B. Lonnerdal, and C. Svanborg. A folding variant of α-lactalbumin with bactericidal activity against Streptococcus pneumoniae. *Mol. Microbiol.* **35**, 589–600 (2000).

[21] M. Svensson, A. Håkansson, A. K. Mossberg, S. Linse, and C. Svanborg. Conversion of α-lactalbumin to a protein inducing apoptosis. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 4221–4226 (2000).

[22] C. Düringer, A. Hamiche, L. Gustafsson, H. Kimura, and C. Svanborg. HAMLET interacts with histones and chromatin in tumor cell nuclei. *J. Biol. Chem.* **278**, 42131–42135 (2003).

[23] D. A. Dolgikh, R. I. Gilmanshin, E. V. Brazhnikov, V. E. Bychkova, G. V. Semisotnov, S. Y. Venyaminov, and O. B. Ptitsyn. α-lactalbumin: compact state with fluctuating tertiary structure? *FEBS Lett.* **136**, 311–315 (1981).

[24] D. P. Kharakoz and V. E. Bychkova. Molten globule of human α-lactalbumin: hydration, density, and compressibility of the interior. *Biochemistry* **36**, 1882–1890 (1997).

[25] M. Kataoka, K. Kuwajima, F. Tokunaga, and Y. Goto. Structural characterization of the molten globule of α-lactalbumin by solution X-ray scattering. *Protein Sci.* **6**, 422–430 (1997).

[26] E. E. Gussakovsky and E. Haas. Two steps in the transition between the native and acid states of bovine α-lactalbumin detected by circular polarization of luminescence: evidence for a premolten globule state? *Protein Sci.* **4**, 2319–2326 (1995).

[27] A. T. Alexandrescu, P. A. Evans, M. Pitkeathly, J. Baum, and C. M. Dobson. Structure and Dynamics of the Acid-Denatured Molten Globule State of Alpha-Lactalbumin - a 2-Dimensional Nmr-Study. *Biochemistry* **32**, 1707–1718 (1993).

[28] S. Bañuelos and A. Muga. Binding of Molten Globule-Like Conformations to Lipid Bilayers - Structure of Native and Partially Folded α-Lactalbumin Bound to Model Membranes. *J. Biol. Chem.* **270**, 29910–29915 (1995).

[29] A. V. Agasøster, . Halskau, E. Fuglebakk, N. A. Frøystein, A. Muga, H. Holmsen, and A. Martínez. The interaction of peripheral proteins and membranes studied with α-lactalbumin and phospholipid bilayers of various compositions. *J. Biol. Chem.* **278**, 21790–21797 (2003).

[30] A. T. Alexandrescu, R. W. Broadhurst, C. Wormald, C. L. Chyan, J. Baum, and C. M. Dobson. ^{1}H-NMR Assignments and Local Environments of Aromatic Residues in Bovine, Human and Guinea-Pig Variants of α-Lactalbumin. *Eur. J. Biochem.* **210**, 699–709 (1992).

[31] V. Forge, R. T. Wijesinha, J. Balbach, K. Brew, C. V. Robinson, C. Redfield, and C. M. Dobson. Rapid Collapse and Slow Structural Reorganisation During the Refolding of Bovine α-Lactalbumin. *J. Mol. Biol.* **288**, 673–688 (1999).

[32] R. Wijesinha-Bettoni, C. M. Dobson, and C. Redfield. Comparison of the structural and dynamical properties of holo and apo bovine α-lactalbumin by NMR spectroscopy. *J. Mol. Biol.* **307**, 885–898 (2001).

[33] F. Bloch, W. W. Hansen, and M. Packard. Nuclear Induction. *Phys. Rev.* **69**, 127–127 (1946).

[34] F. Bloch. Nuclear Induction. *Phys. Rev.* **70**, 460–474 (1946).

[35] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev.* **69**, 37–38 (1946).

[36] R. R. Ernst and W. A. Anderson. Application of Fourier transform spectroscopy to magnetic resonance. *Rev. Sci. Instr.* **37**, 93–102 (1966).

[37] M. H. Levitt. Spin Dynamics: Basics of Nuclear Magnetic Resonance. *Wiley* (2001).

[38] NMRSim 2.8. *Bruker BioSpin* (1998).

[39] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Comp.* **19**, 297–301 (1965).

[40] M. Karplus. Contact electron-spin coupling of nuclear magnetic moments. *J. Phys. Chem.* **30**, 11–15 (1959).

[41] A. Pardi, M. Billeter, and K. Wuthrich. Calibration of the angular dependence of the amide proton-Ca proton coupling constants, ${}^3J_{\mathrm{H^N H^\alpha}}$, in a globular protein. Use of ${}^3J_{\mathrm{H^N H^\alpha}}$ for identification of helical secondary structure. *J. Mol. Biol.* **180**, 741–751 (1984).

[42] M. Piotto, V. Saudek, and V. Sklenar. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661–665 (1992).

[43] V. Sklenar, M. Piotto, R. Leppik, and V. Saudek. Gradient-Tailored Water Suppression for 1H-15N HSQC Experiments Optimized to Retain Full Sensitivity. *J. Magn. Reson. A* **102**, 241–245 (1993).

[44] M. Liu, X. A. Mao, C. Ye, H. Huang, J. K. Nicholson, and J. C. Lindon. Improved WATERGATE Pulse Sequences for Solvent Suppression in NMR Spectroscopy. *J. Magn. Reson.* **132**, 125–129 (1998).

[45] W. P. Aue, E. Bartholdi, and R. R. Ernst. 2-Dimensional Spectroscopy - Application to Nuclear Magnetic-Resonance. *J. Chem. Phys.* **64**, 2229–2246 (1976).

[46] D. I. Hoult. Critical factors in the design of sensitive high resolution NMR spectrometers. *Proc. R. Soc. Lond., Ser. A* **344**, 311–340 (1975).

[47] P. C. M. Vanzijl, M. O. Johnson, S. Mori, and R. E. Hurd. Magic-Angle-Gradient Double-Quantum-Filtered COSY. *J. Magn. Reson. A* **113**, 265–270 (1995).

[48] L. Braunschweiler and R. R. Ernst. Coherence Transfer by Isotropic Mixing - Application to Proton Correlation Spectroscopy. *J. Magn. Reson.* **53**, 521–528 (1983).

[49] A. J. Shaka, C. J. Lee, and A. Pines. Iterative Schemes for Bilinear Operators - Application to Spin Decoupling. *J. Magn. Reson.* **77**, 274–293 (1988).

[50] V. V. Krishnamurthy. Application of Semi-Selective Excitation Sculpting for Homonuclear Decoupling During Evolution in Mulit-Dimensional NMR. *Magn. Reson. Chem.* **35**, 9–12 (1997).

[51] J. Jeener, B. H. Meier, P. Bachmann, and R. R. Ernst. Investigation of exchange process by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546–4563 (1979).

[52] N. A. Frøystein. *Sequence-Selective Interactions of DNA and Metal Ions*. Ph.D. thesis, University of Bergen (1993).

[53] G. Lippens, C. Dhalluin, and J.-M. Wieruszeski. Use of a water flip-back pulse in the homonuclear NOESY experiment. *J. Biomol. NMR* **5**, 327–331 (1995).

[54] V. Sklenar. Suppression of Radiation Damping in Multidimensional NMR Experiments Using Magnetic Field Gradients. *J. Magn. Reson. A* **114**, 132–135 (1995).

[55] G. Bodenhausen and D. J. Ruben. Natural Abundance N-15 NMR by Enhanced Heteronuclear Spectroscopy. *Chem. Phys. Lett.* **69**, 185–189 (1980).

[56] N. A. Frøystein. K304: Practical NMR Spectroscopy - Lecture notes (2001).

[57] S. Mori, C. Abeygunawardana, M. O. Johnson, and P. C. van Zijl. Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J. Magn. Reson. B* **108**, 94–98 (1995).

[58] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, and N. J. Skelton. *Protein NMR Spectroscopy: Principles and Practice*. Academic Press, Inc. (1996).

[59] F. Delaglio, Z. Wu, and A. Bax. Measurement of Homonuclear Proton Couplings from Regular 2D COSY Spectra. *J. Magn. Reson.* **149**, 276–281 (2001).

[60] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

[61] D. S. Wishart, B. D. Sykes, and F. M. Richards. Relationship between Nuclear-Magnetic-Resonance Chemical-Shift and Protein Secondary Structure. *J. Mol. Biol.* **222**, 311–333 (1991).

[62] D. S. Wishart, B. D. Sykes, and F. M. Richards. The Chemical Shift Index: A Fast and Simple Method for the Assignment of Protein Secondary Structure through NMR Spectroscopy. *Biochemistry* **31**, 1647–1651 (1992).

[63] J. N. S. Evans. Biomolecular NMR Spectroscopy. *Oxford University Press* (1995).

[64] J. P. Linge, S. I. O'Donoghue, and M. Nilges. Automated assignment of ambiguous nuclear overhauser effects with ARIA. *Methods Enzymol.* **339**, 71–90 (2001).

[65] M. Nilges. A Calculation Strategy for the Structure Determination of Symmetrical Dimers by ^{1}H-NMR. *Proteins: Struct., Funct., Genet.* **17**, 297–309 (1993).

[66] M. Nilges. Calculation of Protein Structures with Ambiguous Distance Restraints - Automated Assignment of Ambiguous Noe Crosspeaks and Disulfide Connectivities. *J. Mol. Biol.* **245**, 645–660 (1995).

[67] A. T. Brünger et al. Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Cryst.* **D54**, 905–921 (1998).

[68] M. Nilges and S. I. O'Donoghue. Ambiguous NOEs and automated NOE assignment. *Prog. Nucl. Magn. Reson. Spectrosc.* **32**, 107–139 (1998).

[69] D. S. Wishart and B. D. Sykes. The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR* **4**, 171–180 (1994).

[70] B. Kuhlman, D. L. Luisi, P. Young, and D. P. Raleigh. pK_a values and the pH dependent stability of the N-terminal domain of L9 as probes of electrostatic interactions in the denatured state. Differentiation between local and nonlocal interactions. *Biochemistry* **38**, 4896–4903 (1999).

[71] H. A. Chen, M. Pfuhl, M. S. B. McAlister, and P. C. Driscoll. Determination of pK_a values of carboxyl groups in the N-terminal domain of rat CD2: Anomalous pK_a of a glutamate on the ligand-binding surface. *Biochemistry* **39**, 6814–6824 (2000).

[72] Y. Oda, T. Yamazaki, K. Nagayama, S. Kanaya, Y. Kuroda, and H. Nakamura. Individual Ionization-Constants of All the Carboxyl Groups in Ribonuclease Hi from Escherichia-Coli Determined by NMR. *Biochemistry* **33**, 5275–5284 (1994).

[73] B. M. Fisher, L. W. Schultz, and R. T. Raines. Coulombic effects of remote subsites on the active site of ribonuclease A. *Biochemistry* **37**, 17386–17401 (1998).

[74] J. M. Perez-Canadillas, R. Campos-Olivas, J. Lacadena, A. M. del Pozo, J. G. Gavilanes, J. Santoro, M. Rico, and M. Bruix. Characterization of pK_a values and titration shifts in the cytotoxic ribonuclease α-sarcin by NMR. Relationship between electrostatic interactions, structure, and catalytic function. *Biochemistry* **37**, 15865–15876 (1998).

[75] J. L. Markley. Observation of Histidine Residues in Proteins by Means of Nuclear Magnetic-Resonance Spectroscopy. *Acc. Chem. Res.* **8**, 70–80 (1975).

[76] M. Sundd, N. Iverson, B. Ibarra-Molero, J. M. Sanchez-Ruiz, and A. D. Robertson. Electrostatic interactions in ubiquitin: stabilization of carboxylates by lysine amino groups. *Biochemistry* **41**, 7586–7596 (2002).

[77] T. D. Goddard and D. G. Kneller. SPARKY 3. *University of California, San Francisco* (2002).

[78] R. Koradi, M. Billeter, M. Engeli, P. Guntert, and K. Wuthrich. Automated Peak Picking and Peak Integration in Macromolecular NMR Spectra Using AUTOPSY. *J. Magn. Reson.* **135**, 288–297 (1998).

[79] N. C. J. Strynadka, M. Cherney, A. R. Sielecki, M. X. Li, L. B. Smillie, and M. N. G. James. Structural details of a calcium-induced molecular switch: X-ray crystallographic analysis of the calcium-saturated N-terminal domain of troponin C at 1.75 Å resolution. *J. Mol. Biol.* **273**, 238–255 (1997).

[80] D. R. Bauer, S. J. Opella, D. J. Nelson, and R. Pecora. Depolarized Light-Scattering and Carbon Nuclear-Resonance Measurements of Isotropic Rotational Correlation Time of Muscle Calcium-Binding Protein. *J. Am. Chem. Soc.* **97**, 2580–2582 (1975).

[81] DataFit 6.1.10. *Oakdale Engineering* (1999).

[82] B. Ibarra-Molero, V. V. Loladze, G. I. Makhatadze, and J. M. Sanchez-Ruiz. Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry* **38**, 8138–8149 (1999).

[83] M. Williamson and M. Rafaee. SHIFTCALC: Chemical shift calculations. *NMR Group, MBB, The University of Sheffield* (2003).

[84] K. Harata and M. Muraki. X-ray structural evidence for a local helix-loop transition in α-lactalbumin. *J. Biol. Chem.* **267**, 1419 (1992).

[85] A. C. Pike, K. Brew, and K. R. Acharya. Crystal structures of guinea-pig, goat and bovine α-lactalbumin highlight the enhanced conformational flexibility of regions that are significant for its action in lactose synthase. *Structure* **4**, 691–703 (1996).

[86] V. A. Malinovskii, J. Tian, J. A. Grobler, and K. Brew. Functional site in α-lactalbumin encompasses a region corresponding to a subsite in lysozyme and parts of two adjacent flexible substructures. *Biochemistry* **35**, 9710–9715 (1996).

[87] L. H. Greene, J. A. Grobler, V. A. Malinovskii, J. Tian, K. R. Acharya, and K. Brew. Stability, activity and flexibility in α-lactalbumin. *Protein Eng.* **12**, 581–587 (1999).

[88] B. Ramakrishnan and P. K. Qasba. Crystal Structure of Lactose Synthase Reveals a Large Conformational Change in its Catalytic Component, the β1,4-Galactosyltransferase I,. *J. Mol. Biol.* **310**, 205–218 (2001).

[89] H. Schwalbe, S. B. Grimshaw, A. Spencer, M. Buck, J. Boyd, C. M. Dobson, C. Redfield, and L. J. Smith. A refined solution structure of hen lysozyme determined using residual dipolar coupling data. *Protein Sci.* **10**, 677–688 (2001).

[90] L. J. Smith, M. J. Sutcliffe, C. Redfield, and C. M. Dobson. Structure of Hen Lysozyme in Solution. *J. Mol. Biol.* **229**, 930–944 (1993).

[91] M. Ramanadham, L. C. Sieker, and L. H. Jensen. Refinement of triclinic lysozyme: II. The method of stereochemically restrained least squares. *Acta Crystallogr. B.* **46**, 63–69 (1990).

[92] N. Ishikawa, T. Chiba, L. Chen, A. Shimizu, M. Ikeguchi, and S. Sugai. Remarkable destabilization of recombinant α-lactalbumin by an extraneous N-terminal methionyl residue. *Protein Eng.* **11**, 333–335 (1998).

Appendix

A.1 Pulse programs

A.1.1 1D ^{1}H

```
;zg
;avance-version
;1D sequence

#include <Avance.incl>

1 ze
2 d1
  p1 ph1
  go=2 ph31
  wr #0
exit

ph1=0 2 2 0 1 3 3 1
ph31=0 2 2 0 1 3 3 1

;pl1 : f1 channel - power level for pulse (default)
;p1 : f1 channel -  high power pulse
;d1 : relaxation delay; 1-5 * T1
```

A.1.2 WATERGATE

```
;dpfgsew5
;avance-version
;1D sequence
;water suppression using w5 pulse sequence with gradients
;and excitation sculpting
;
;M. Liu, X. Mao, C. Ye, H. Huang, J.K. Nicholson and J.C. Lindon,
;  J. Magn. Reson., 132, 125-129 (1998)
;T.-L. Hwang & A.J. Shaka, J. Magn. Reson.,
;  Series A 112, 275-279 (1995).

#include <Avance.incl>
#include <Grad.incl>

1 ze
2 d1 pl1:f1
  p1 ph1
  50u UNBLKGRAD
  GRADIENT(cnst21)
  d16 pl18:f1
  p28*0.0867 ph3
```

```
  d19*2
  p28*0.2056 ph3
  d19*2
  p28*0.4133 ph3
  d19*2
  p28*0.7778 ph3
  d19*2
  p28*1.4911 ph3
  d19*2
  p28*1.4911 ph4
  d19*2
  p28*0.7778 ph4
  d19*2
  p28*0.4133 ph4
  d19*2
  p28*0.2056 ph4
  d19*2
  p0*0.0867 ph4
  50u
  GRADIENT(cnst22)
  d16
  50u
  GRADIENT(cnst23)
  d16
  p28*0.0867 ph3
  d19*2
  p28*0.2056 ph3
  d19*2
  p28*0.4133 ph3
  d19*2
  p28*0.7778 ph3
  d19*2
  p28*1.4911 ph3
  d19*2
  p28*1.4911 ph4
  d19*2
  p28*0.7778 ph4
  d19*2
  p28*0.4133 ph4
  d19*2
  p28*0.2056 ph4
  d19*2
  p0*0.0867 ph4
  46u
  GRADIENT(cnst24)
  d16
  4u BLKGRAD
  go=2 ph31
  wr #0
exit

ph1=0 2
ph3=0
ph4=2
ph31=0 2

;pl1 : f1 channel - power level for pulse (default)
;pl18: f1 channel - power level for w5-pulse
;p0 : f1 channel -  90 degree pulse at pl18, use for fine adjustment
;p1 : f1 channel -  90 degree high power pulse
;p16: homospoil/gradient pulse
;p28: f1 channel -  90 degree pulse at pl18
```

```
;d1 : relaxation delay; 1-5 * T1
;d16: delay for homospoil/gradient recovery
;d19: delay for binomial water suppression
;     d19 = (1/(2*d)), d = distance of next null (in Hz)
;NS: 8 * n
;DS: 4

;use gradient program (GRDPROG) : 4sine
;use gradient ratio: cnst21 : cnst22 : cnst23 : cnst24
;     40 :     40 :      7 :      7
```

A.1.3 COSY with presaturation

```
;cosyprst
;avance-version
;2D homonuclear shift correlation
;with presaturation during relaxation delay
;phase sensitive using States-TPPI method

#include <Avance.incl>

"d0=3u"
"d11=30m"
"d12=20u"
"d13=3u"
"l3=(td1/2)"

1 ze
2 d11
  3m
3 d11
4 d12 pl9:f1
  d1 cw:f1 ph29
  d13 do:f1
  d12 pl1:f1
  p1 ph1
  d0
  p0 ph2
  go=2 ph31
  d11 wr #0 if #0 ip1 zd
  3m ip29
  lo to 3 times 2
  d11 id0
  lo to 4 times l3
exit

ph1=0 2 2 0 1 3 3 1
ph2=0 2 0 2 1 3 1 3
ph29=0
ph31=0 2 2 0 1 3 3 1

;pl1 : f1 channel - power level for pulse (default)
;pl9 : f1 channel - power level for presaturation
;p0 : f1 channel -  20 to 90 degree high power pulse
;p1 : f1 channel -  90 degree high power pulse
;d0 : incremented delay (2D)                          [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O                              [30 msec]
;d12: delay for power switching                       [20 usec]
;d13: short delay                                     [3 usec]
;l3: loop for phase sensitive 2D using States-TPPI method: l3 = td1/2
;in0: 1/(1 * SW) = 2 * DW
```

```
;nd0: 1
;NS: 4 * n
;DS: 16
;td1: number of experiments
;MC2: States-TPPI
```

A.1.4 1D DQ filter

```
;mqsgs1d
;avance-version
;1D multiple quantum filter
;using gradient pulses for selection
;with multiple quantum filter according to gradient ratio
;use as setup for pulseprogram 'cosygsmftp'
;   or for magic angle gradient calibration

#include <Avance.incl>
#include <Grad.incl>

"d13=3u"
"d20=p16+d16"

1 ze
2 d1
3 p1 ph1
  d4
  p2 ph4
  d4
  p1 ph2
  d13
  d20 UNBLKGRAD
  p2 ph4
  d13
  GRADIENT(cnst21)
  d16
  p1 ph3
  d20
  p2 ph4
  d13
  GRADIENT(cnst22)
  d16
  4u BLKGRAD
  go=2 ph31
  wr #0
exit

ph1=0 2
ph2=0
ph3=0
ph4=1
ph31=0 2

;pl1 : f1 channel - power level for pulse (default)
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p16: homospoil/gradient pulse
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J)
;d13: short delay                                   [3 usec]
;d16: delay for homospoil/gradient recovery
;d20: p16 + d16
;in0: 1/(2 * SW) = DW
```

```
;nd0: 2
;NS: 1 * n
;DS: 16

;use gradient program (GRDPROG) :
;   for z-gradients only:                  2sine
;   use gradient ratio:
;            cnst21 : cnst22
;              10 :     20         for double quantum filter (dqf)
;           10 :     30         for triple quantum filter (tqf)
;   for magic angle gradients:             2sinexyz
;   use gradient ratio:
;            cnst21 : cnst22 : cnst23 : cnst24 : cnst25 : cnst26
;           k*10 :   k*10 :     10 :   k*20     k*20       20  (dqf)
;        k*10 :   k*10 :     10 :   k*30     k*30       30  (tqf)
```

A.1.5 gs-DQF-COSY

```
;cosygpmftp
;avance-version
;2D homonuclear shift correlation
;phase sensitive using TPPI
;using gradient pulses for selection
;with multiple quantum filter according to gradient ratio
;use pulseprogram 'mqsgp1d' for setup

#include <Avance.incl>
#include <Grad.incl>

"p2=p1*2"
"d0=3u"
"d13=3u"
"d20=p16+d16+d13"

1 ze
2 d1
3 p1 ph1
  d0
  p1 ph2
  d20 UNBLKGRAD
  p2 ph4
  d13
  p16:gp1
  d16
  p1 ph3
  d20
  p2 ph4
  d13
  p16:gp2
  d16
  4u BLKGRAD
  go=2 ph31
  d1 wr #0 if #0 id0 ip1 zd
  lo to 3 times td1
exit

ph1=0 2
ph2=0
ph3=0
ph4=0
ph31=0 2
```

```
;pl1 : f1 channel - power level for pulse (default)
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p16: homospoil/gradient pulse
;d0 : incremented delay (2D)                          [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d13: short delay                                     [3 usec]
;d16: delay for homospoil/gradient recovery
;d20: p16 + d16 + d13
;in0: 1/(2 * SW) = DW
;nd0: 2
;NS: 1 * n
;DS: 16
;td1: number of experiments
;MC2: TPPI

;use gradient ratio: gp 1 : gp 1
;   10 :   20   for double quantum filter
;   10 :   30   for triple quantum filter

;for z-only gradients:
;gpz1: 10%
;gpz2: 20% for DQF, 30% for TQF

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100
```

A.1.6 TOCSY

```
;dipsi2stdpfgsew5
;avance-version
;homonuclear Hartman-Hahn transfer using DIPSI2 sequence for mixing
;phase sensitive using States-TPPI
;
;water suppression using w5 pulse sequence with gradients
;and excitation sculpting
;M. Liu, X. Mao, C. Ye, H. Huang, J.K. Nicholson and J.C. Lindon,
;  J. Magn. Reson., 132, 125-129 (1998)
;T.-L. Hwang & A.J. Shaka, J. Magn. Reson.,
;  Series A 112, 275-279 (1995).

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"d0=3u"
"d11=30m"
"d12=20u"
"d20=2m"
"d21=3m"

"FACTOR1=(d9/(p6*115.112))/2+0.5"
"l1=FACTOR1*2"
"l3=(td1/2)"

1 ze
2 d1
3 d11
4 d12 pl1:f1
  p1 ph1
  d0
```

```
  p1 ph2
  d20 pl10:f1
5 p6*3.556 ph23
  p6*4.556 ph25
  p6*3.222 ph23
  p6*3.167 ph25
  p6*0.333 ph23
  p6*2.722 ph25
  p6*4.167 ph23
  p6*2.944 ph25
  p6*4.111 ph23
  p6*3.556 ph25
  p6*4.556 ph23
  p6*3.222 ph25
  p6*3.167 ph23
  p6*0.333 ph25
  p6*2.722 ph23
  p6*4.167 ph25
  p6*2.944 ph23
  p6*4.111 ph25
  p6*3.556 ph25
  p6*4.556 ph23
  p6*3.222 ph25
  p6*3.167 ph23
  p6*0.333 ph25
  p6*2.722 ph23
  p6*4.167 ph25
  p6*2.944 ph23
  p6*4.111 ph25
  p6*3.556 ph23
  p6*4.556 ph25
  p6*3.222 ph23
  p6*3.167 ph25
  p6*0.333 ph23
  p6*2.722 ph25
  p6*4.167 ph23
  p6*2.944 ph25
  p6*4.111 ph23
  lo to 5 times l1
  d21 pl1:f1
  p1 ph3
  50u UNBLKGRAD
  GRADIENT(cnst21)
  d16 pl18:f1
  p28*0.0867 ph4
  d19*2
  p28*0.2056 ph4
  d19*2
  p28*0.4133 ph4
  d19*2
  p28*0.7778 ph4
  d19*2
  p28*1.4911 ph4
  d19*2
  p28*1.4911 ph5
  d19*2
  p28*0.7778 ph5
  d19*2
  p28*0.4133 ph5
  d19*2
  p28*0.2056 ph5
  d19*2
```

```
  p0*0.0867 ph5
  50u
  GRADIENT(cnst22)
  d16
  50u
  GRADIENT(cnst23)
  d16
  p28*0.0867 ph4
  d19*2
  p28*0.2056 ph4
  d19*2
  p28*0.4133 ph4
  d19*2
  p28*0.7778 ph4
  d19*2
  p28*1.4911 ph4
  d19*2
  p28*1.4911 ph5
  d19*2
  p28*0.7778 ph5
  d19*2
  p28*0.4133 ph5
  d19*2
  p28*0.2056 ph5
  d19*2
  p0*0.0867 ph5
  46u
  GRADIENT(cnst24)
  d16
  4u BLKGRAD
  go=2 ph31
  d1 wr #0 if #0 ip1 zd
  lo to 3 times 2
  d11 id0
  lo to 4 times l3
exit

ph1=0 2
ph2=0 0 0 0 2 2 2 2
ph3=0 0 2 2
ph4=0
ph5=2
ph23=3
ph25=1
ph31=0 2 2 0 2 0 0 2

;pl1 : f1 channel - power level for pulse (default)
;pl10: f1 channel - power level for TOCSY-spinlock
;pl18: f1 channel - power level for w5 sequence
;p0 : f1 channel -  90 degree pulse at pl18, use for fine adjustment
;p1 : f1 channel -  90 degree high power pulse
;p6 : f1 channel -  90 degree low power pulse
;p28: f1 channel -  90 degree pulse at pl18
;d0 : incremented delay (2D)                          [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d9 : TOCSY mixing time
;d11: delay for disk I/O                              [30 msec]
;d12: delay for power switching                       [20 usec]
;d19: delay for binomial water suppression
;     d19 = (1/(2*d)), d = distance of next null (in Hz)
;d20: first z-filter delay                            [2 msec]
;d21: second z-filter delay                           [3 msec]
```

```
;l1: loop for DIPSI cycle: ((p6*115.112) * l1) = mixing time
;l3: loop for phase sensitive 2D using States-TPPI method: l3 = td1/2
;in0: 1/(1 * SW) = 2 * DW
;nd0: 1
;NS: 8 * n
;DS: 16
;td1: number of experiments
;MC2: States-TPPI

;use gradient program (GRDPROG) : 4sine
;use gradient ratio: cnst21 : cnst22 : cnst23 : cnst24
;     40 :     40 :      7 :      7
```

A.1.7 BASHD-TOCSY

```
;dipsi2bashdxx
;avance-version
;homonuclear Hartman-Hahn transfer using DIPSI2 sequence for mixing
;phase sensitive using States-TPPI
;Double PFG spin-echo for band selective excitation and F1 homonuclear decoupling
;
;V. V. Krishnamurthy, J. Magn. Reson., Series B 113, 46-52 (1996).
;V. V. Krishnamurthy, Magn. Reson. Chem., 35, 9-12 (1997).
;A. Kaerner and D. L. Rabenstein, Magn. Reson. Chem., 36, 601-617 (1998).
;
;water suppression using w5 pulse sequence with gradients
;and excitation sculpting
;
;M. Liu, X. Mao, C. Ye, H. Huang, J.K. Nicholson and J.C. Lindon,
;  J. Magn. Reson., 132, 125-129 (1998)
;T.-L. Hwang & A.J. Shaka, J. Magn. Reson.,
;  Series A 112, 275-279 (1995).

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"d0=3u"
;;"d0=0.5*(in0-p1*4/3.14159265-p2)"
"d11=30m"
"d12=20u"
"d20=2m"
"d21=3m"
"FACTOR1=(d9/(p6*115.112))/2+0.5"
"l1=FACTOR1*2"
"l3=(td1/2)"

1 ze
2 d1
3 d11
4 d12 pl1:f1
  p1 ph1
  d0
  p2 ph22
  50u UNBLKGRAD
  GRADIENT(cnst21)
  d16 pl0:f1
  (p11:sp1 ph12):f1
  50u
  GRADIENT(cnst22)
  d16
  d0
```

```
  50u
  GRADIENT(cnst23)
  d16 pl0:f1
  (p11:sp1 ph13):f1
  46u
  GRADIENT(cnst24)
  d16
  4u BLKGRAD
  p1 ph2
  d20 pl10:f1
5 p6*3.556 ph23
  p6*4.556 ph25
  p6*3.222 ph23
  p6*3.167 ph25
  p6*0.333 ph23
  p6*2.722 ph25
  p6*4.167 ph23
  p6*2.944 ph25
  p6*4.111 ph23
  p6*3.556 ph25
  p6*4.556 ph23
  p6*3.222 ph25
  p6*3.167 ph23
  p6*0.333 ph25
  p6*2.722 ph23
  p6*4.167 ph25
  p6*2.944 ph23
  p6*4.111 ph25
  p6*3.556 ph25
  p6*4.556 ph23
  p6*3.222 ph25
  p6*3.167 ph23
  p6*0.333 ph25
  p6*2.722 ph23
  p6*4.167 ph25
  p6*2.944 ph23
  p6*4.111 ph25
  p6*3.556 ph23
  p6*4.556 ph25
  p6*3.222 ph23
  p6*3.167 ph25
  p6*0.333 ph23
  p6*2.722 ph25
  p6*4.167 ph23
  p6*2.944 ph25
  p6*4.111 ph23
  lo to 5 times l1
  d21 pl1:f1
  p1 ph3
  50u UNBLKGRAD
  GRADIENT(cnst21)
  d16 pl18:f1
  p28*0.0867 ph4
  d19*2
  p28*0.2056 ph4
  d19*2
  p28*0.4133 ph4
  d19*2
  p28*0.7778 ph4
  d19*2
  p28*1.4911 ph4
  d19*2
```

```
  p28*1.4911 ph5
  d19*2
  p28*0.7778 ph5
  d19*2
  p28*0.4133 ph5
  d19*2
  p28*0.2056 ph5
  d19*2
  p0*0.0867 ph5
  50u
  GRADIENT(cnst22)
  d16
  50u
  GRADIENT(cnst23)
  d16
  p28*0.0867 ph4
  d19*2
  p28*0.2056 ph4
  d19*2
  p28*0.4133 ph4
  d19*2
  p28*0.7778 ph4
  d19*2
  p28*1.4911 ph4
  d19*2
  p28*1.4911 ph5
  d19*2
  p28*0.7778 ph5
  d19*2
  p28*0.4133 ph5
  d19*2
  p28*0.2056 ph5
  d19*2
  p0*0.0867 ph5
  46u
  GRADIENT(cnst24)
  d16
  4u BLKGRAD
  go=2 ph31
  d1 wr #0 if #0 ip1 zd
  lo to 3 times 2
  d11 id0
  lo to 4 times l3
exit

ph1=0 2
ph2=0 0 0 0 2 2 2 2
ph3=0 0 2 2
ph4=0
ph5=2
ph12=0 2
ph13=2 0
ph22=0 2
ph23=3
ph25=1
ph31=0 2 2 0 2 0 0 2

;pl0 : 120dB
;pl1 : f1 channel - power level for pulse (default)
;pl10: f1 channel - power level for TOCSY-spinlock
;pl18: f1 channel - power level for w5 sequence
;sp1 : f1 channel - shaped pulse
```

```
;p0 : f1 channel -  90 degree pulse at pl18, use for fine adjustment
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel -  180 degree high power pulse
;p6 : f1 channel -  90 degree low power pulse
;p11: f1 channel -  180 degree shaped pulse
;p16: homospoil/gradient pulse
;p28: f1 channel -  90 degree pulse at pl18
;d0 : incremented delay (2D)                          [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d9 : TOCSY mixing time
;d11: delay for disk I/O                              [30 msec]
;d12: delay for power switching                       [20 usec]
;d16: delay for homospoil/gradient recovery
;d19: delay for binomial water suppression
;     d19 = (1/(2*d)), d = distance of next null (in Hz)
;d20: first z-filter delay                            [2 msec]
;d21: second z-filter delay                           [3 msec]
;l1: loop for DIPSI cycle: ((p6*115.112) * l1) = mixing time
;l3: loop for phase sensitive 2D using States-TPPI method: l3 = td1/2
;in0: 1/(2 * SW) = DW
;nd0: 2
;NS: 8 * n
;DS: 16
;td1: number of experiments
;MC2: States-TPPI

;use gradient program (GRDPROG) : 8sine
;use gradient ratio: cnst21 : cnst22 : cnst23 : cnst24 : cnst25 : cnst26 : cnst27 : const28
;     3  :      3 :     11 :     11 :     20 :     20 :      7 :       7
;use shaped pulse program (spnam1) : GaussCascadeQ3
```

A.1.8 NOESY

```
;noesyfprsdpw5
;avance-version
;2D homonuclear correlation via dipolar coupling
;dipolar coupling may be due to noe or chemical exchange.
;phase sensitive using States-TPPI method
;water suppression using w5 pulse sequence with gradients
;and excitation sculpting
;using flip-back pulse
;with radiation damping suppression using gradients in t1
;
;M. Piotto, V. Saudek & V. Sklenar, J. Biomol. NMR 2,
;   661 - 666 (1992)
;V. Sklenar, M. Piotto, R. Leppik & V. Saudek, J. Magn. Reson. A102,
;   241 -245 (1993)
;G. Lippens, C. Dhalluin & J.-M. Wieruszeski, J. Biomol. NMR 5,
;   327-331 (1995)
;V. Sklenar, J. Magn. Reson. A114, 132-135 (1995)
;M. Liu, X. Mao, C. Ye, H. Huang, J. K. Nicholson & J. C. Lindon, J. Magn.
;   Reson. 132, 125-129 (1998)
;T.-L. Hwang & A.J. Shaka, J. Magn. Reson.,
;   Series A 112, 275-279 (1995).

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

;;"d0=3u"
"d11=30m"
"d12=20u"
```

```
"DELTA=d8-p16-d16-p11-d12-54u"
"d0=0.5*(in0-p1*4/3.14159265-6u)"
"l3=(td1/2)"

1 ze
2 d1
3 d11
4 d12 pl1:f1
  p1 ph1
  2u:ngrad
  d0
  2u:ngrad
  d0
  2u:ngrad
  p1 ph2
  DELTA
  50u UNBLKGRAD
  GRADIENT(cnst22)
  d16 pl0:f1
  (p11:sp1 ph3:r):f1
  4u
  d12 pl1:f1
  p1 ph4
  8u pl18:f1
  GRADIENT(cnst23)
  d16
  p28*0.0867 ph5
  d19*2
  p28*0.2056 ph5
  d19*2
  p28*0.4133 ph5
  d19*2
  p28*0.7778 ph5
  d19*2
  p28*1.4911 ph5
  d19*2
  p28*1.4911 ph6
  d19*2
  p28*0.7778 ph6
  d19*2
  p28*0.4133 ph6
  d19*2
  p28*0.2056 ph6
  d19*2
  p0*0.0867 ph6
  50u
  GRADIENT(cnst23)
  d16
  50u
  GRADIENT(cnst24)
  d16
  p28*0.0867 ph5
  d19*2
  p28*0.2056 ph5
  d19*2
  p28*0.4133 ph5
  d19*2
  p28*0.7778 ph5
  d19*2
  p28*1.4911 ph5
  d19*2
  p28*1.4911 ph6
```

```
  d19*2
  p28*0.7778 ph6
  d19*2
  p28*0.4133 ph6
  d19*2
  p28*0.2056 ph6
  d19*2
  p0*0.0867 ph6
  46u
  GRADIENT(cnst24)
  d16
  4u BLKGRAD
  go=2 ph31
  d1 wr #0 if #0 ip1 zd
  lo to 3 times 2
  d11 id0
  lo to 4 times l3
exit

ph1=0 2
ph2=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph3=2 2 0 0 3 3 1 1
ph4=0 0 2 2 1 1 3 3
ph5=0
ph6=2
ph31=0 2 2 0 1 3 3 1 2 0 0 2 3 1 1 3

;pl0 : 120dB
;pl1 : f1 channel - power level for pulse (default)
;pl18: f1 channel - power level for w5 pulse
;sp1: f1 channel - shaped pulse  90 degree
;p0 : f1 channel -  90 degree pulse at pl18, use for fine adjustment
;p1 : f1 channel -  90 degree high power pulse
;p11: f1 channel -  90 degree shaped pulse
;p16: homospoil/gradient pulse
;p28: f1 channel -  90 degree pulse at pl18
;d0 : incremented delay (2D)
;d1 : relaxation delay; 1-5 * T1
;d8 : mixing time
;d11: delay for disk I/O                              [30 msec]
;d12: delay for power switching                       [20 usec]
;d16: delay for homospoil/gradient recovery
;d19: delay for binomial water suppression
;     d19 = (1/(2*d)), d = distance of next null (in Hz)
;l3: loop for phase sensitive 2D using States-TPPI method: l3 = td1/2
;in0: 1/(2 * SW) = DW
;nd0: 2
;NS: 8 * n
;DS: 16
;td1: number of experiments
;MC2: States-TPPI

;use gradient program (GRDPROG) : Noesyfprsdp
;use gradient ratio: cnst21 : cnst22 : cnst23 : cnst24
;      2 :     50 :     34 :     22
```

A.1.9 ^{1}H-^{15}N HSQC

```
;invif3gsst19
;avance-version
;2D H-1/X correlation via double inept transfer
;phase sensitive using States-TPPI method
```

```
;with decoupling during acquisition
;using f3 - channel
;water suppression using 3-9-19 pulse sequence with gradients
;G. Bodenhausen & D.J. Ruben, Chem. Phys. Lett. 69, 185 (1980)
;M. Piotto, V. Saudek & V. Sklenar, J. Biomol. NMR 2, 661 - 666 (1992)
;V. Sklenar, M. Piotto, R. Leppik & V. Saudek, J. Magn. Reson.,
;   Series A 102, 241 -245 (1993)

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"p22=p21*2"
"d0=3u"
"d11=30m"
"d12=20u"
"d13=4u"
"d26=1s/(cnst4*4)"
"DELTA=d19-p22/2"
"DELTA1=d26-p16-d16-p29*2.385-d19*5+p22/2"
"DELTA2=d26-p16-d16-p29*2.154-p0*0.231-d19*5+p22/2-8u"
"CEN_HN2=(p22-p2)/2"
"l3=(td1/2)"

1 ze
  d11 pl16:f3
2 d1 do:f3
  3m
3 d11
4 d12 pl1:f1
  (p1 ph1)
  d26 pl3:f3
  (CEN_HN2 p2 ph1) (p22 ph6):f3
  d26 UNBLKGRAD
  (p28 ph1)
  d13
  (p1 ph2)
  3u
  GRADIENT(cnst21)
  d16
  (p21 ph3):f3
  d0
  (p2 ph5)
  d0
  (p21 ph4):f3
  3u
  GRADIENT(cnst22)
  d16
  (p1 ph1)
  DELTA1
  GRADIENT(cnst23)
  d16 pl18:f1
  p29*0.231 ph7
  d19*2
  p29*0.692 ph7
  d19*2
  p29*1.462 ph7
  DELTA
  (p22 ph1):f3
  DELTA
  p29*1.462 ph8
```

```
  d19*2
  p29*0.692 ph8
  d19*2
  p0*0.231 ph8
  4u
  GRADIENT(cnst24)
  d16
  4u BLKGRAD
  DELTA2 pl16:f3
  go=2 ph31 cpd3:f3
  d1 do:f3 wr #0 if #0 ip3 zd
  3m ip6
  lo to 3 times 2
  d11 id0
  lo to 4 times l3
exit

ph1=0
ph2=1
ph3=0 2
ph4=0 0 0 0 2 2 2 2
ph5=0 0 2 2
ph6=0
ph7=0
ph8=2
ph31=0 2 0 2 2 0 2 0

;pl1 : f1 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl16: f3 channel - power level for CPD/BB decoupling
;pl18: f1 channel - power level for 3-9-19-pulse (watergate)
;p0 : f1 channel -  90 degree pulse at pl18, use for fine adjustment
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p16: homospoil/gradient pulse
;p21: f3 channel -  90 degree high power pulse
;p22: f3 channel - 180 degree high power pulse
;p28: f1 channel - trim pulse                          [1 msec]
;p29: f1 channel -  90 degree pulse at pl18
;d0 : incremented delay (2D)                          [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O                              [30 msec]
;d12: delay for power switching                       [20 usec]
;d13: short delay                                     [4 usec]
;d16: delay for homospoil/gradient recovery
;d19: delay for binomial water suppression
;     d19 = (1/(2*d)), d = distance of next null (in Hz)
;d26 : 1/(4J)YH
;cnst4: = J(YH)
;l3: loop for phase sensitive 2D using States-TPPI method: l3 = td1/2
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;NS: 4 * n
;DS: 16
;td1: number of experiments
;MC2: States-TPPI
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;use gradient program (GRDPROG) : 4sine
;use gradient ratio: cnst21 : cnst22 : cnst23 : cnst24
;     80 :    -50 :     30 :     30
```

A.1.10 ^{1}H-^{15}N FHSQC

```
;fhsqc.nf
;avance-version
;2D H-1/X correlation via double inept transfer
;phase sensitive using States-TPPI method
;with decoupling during acquisition
;S. Mori, C. Abeygunawardana, M. O'Neil-Johnson & P.C.M. van Zijl,
;   J. Magn. Reson. B 108, 94-98 (1995)
;   N.AA.F. 11/7-00: f3 instead of f2

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

;;"p2=p1*2"
;;"p22=p21*2"
;;"d0=3u"
;;"d11=30m"
;;"d12=20u"
;;"d13=4u"
;;"d26=1s/(cnst4*4)"

"DELTA=d19-p22/2"
"DELTA1=d26-p16-d16-p28*3-d19*5+p22/2"
"DELTA2=d26-p16-d16-p28*2-p0-d19*5+p22/2-8u"
"TAU=d26-p16-4u"
"CEN_HN2=(p22-p2)/2"
"l3=(td1/2)"

1 ze
  d11 pl16:f2
2 d1 do:f2
  3m
3 d11
4 d12 pl1:f1
  50u UNBLKGRAD
  (p1 ph1)
  4u
  GRADIENT(cnst21)
  TAU pl2:f2
  (CEN_HN2 p2 ph1) (p22 ph6):f2
  4u
  GRADIENT(cnst21)
  TAU
  (p1 ph2)
  4u
  GRADIENT(cnst22)
  d16
  (p21 ph3):f2
  d0
  (p2 ph5)
  d0
  (p21 ph4):f2
  4u
  GRADIENT(cnst22)
  d16
  (p1 ph7)
  DELTA1
  GRADIENT(cnst23)
  d16 pl18:f1
  p28*0.231 ph8
  d19*2
```

```
  p28*0.692 ph8
  d19*2
  p28*1.462 ph8
  DELTA
  (p22 ph1):f2
  DELTA
  p28*1.462 ph9
  d19*2
  p28*0.692 ph9
  d19*2
  p0*0.231 ph9
  4u
  GRADIENT(cnst23)
  d16
  4u BLKGRAD
  DELTA2 pl16:f2
  go=2 ph31 cpd3:f2
  d1 do:f2 wr #0 if #0 ip3 zd
  3m ip6
  lo to 3 times 2
  d11 id0
  lo to 4 times l3
exit

ph1=0
ph2=1
ph3=0 2
ph4=0 0 0 0 2 2 2 2
ph5=0 0 2 2
ph6=0
ph7=2
ph8=1
ph9=3
ph31=0 2 0 2 2 0 2 0

;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl16: f2 channel - power level for CPD/BB decoupling
;pl18: f1 channel - power level for 3-9-19-pulse (watergate)
;p0 : f1 channel -  90 degree pulse at pl18, use for fine adjustment
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p16: homospoil/gradient pulse
;p21: f2 channel -  90 degree high power pulse
;p22: f2 channel - 180 degree high power pulse
;p28: f1 channel -  90 degree pulse at pl18
;d0 : incremented delay (2D)                          [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O                              [30 msec]
;d12: delay for power switching                       [20 usec]
;d13: short delay                                     [4 usec]
;d16: delay for homospoil/gradient recovery
;d19: delay for binomial water suppression
;     d19 = (1/(2*d)), d = distance of next null (in Hz)
;d26 : 1/(4J(YH))
;cnst4: = J(YH)
;l3: loop for phase sensitive 2D using States-TPPI method: l3 = td1/2
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;NS: 8 * n
;DS: 16
;td1: number of experiments
```

```
;MC2: States-TPPI
;cpd2: decoupling according to sequence defined by cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence

;use gradient program (GRDPROG) :
;   for z only:                          Fhsqcz
;      with gradient ratio:  cnst21 : cnst22 : cnst23
;          50 :     80 :     30
;   for xyz:                             Fhsqcxyz
;      with gradient ratio:
;               cnst21 : cnst22 : cnst23 : cnst24 : cnst25 : cnst26
;                   20 :     40 :     20 :     10 :   k*15 :     15
;      with k = calibration constant from 1D magic angle gradient calibration
```

A.2 Assigned resonances

Table A.2: The assigned resonances of holo-BLA at 314 K and pH 5.6. All values are in ppm.

Resid.	H^N	H^α	^{15}N	H^β/$H^{\beta 2/\beta 3}$	Other resonances
E1					
Q2					
L3					
T4	7.87	4.70	107.9		
K5	8.75	3.56	120.8		
C6	8.57	4.66	112.7	3.14/3.10	
E7	7.65	4.18	121.7	2.32	
V8	8.16	3.18	120.7	1.89	$H^{\gamma 1/\gamma 2}$=0.47/0.78
F9	7.94	4.19	117.8		
R10	7.28	4.12	114.6	2.07	$H^{\gamma 2/\gamma 3}$=1.63/1.92
E11	8.57	4.14	117.1	1.78	$H^{\gamma 2/\gamma 3}$=2.49
L12	7.86	4.20	117.1		
K13	6.89	3.88	119.2	1.93/2.06	$H^{\epsilon 2/\epsilon 3}$=3.07
D14	9.19	4.57	118.7	2.46/2.55	
L15	8.25	4.16	116.7	1.35/1.24	H^{γ}=1.03, $H^{\delta 1/\delta 2}$=0.17/-0.26
K16	7.68	3.73	122.3		
G17	8.67	3.52	118.2		
Y18	8.25	4.17	127.0		$H^{\delta 1/\delta 2}$=7.12, $H^{\epsilon 1/\epsilon 2}$=6.98
G19	8.85	3.90/3.76	118.0		
G20	8.45	4.82/3.32	127.2		
V21	7.05	3.81	122.5	2.26	$H^{\gamma 1/\gamma 2}$=0.01/0.61
S22	8.14	4.60	122.5		
L23	9.52	4.54	120.1		
P24					
E25	7.09	3.80	113.3		
W26	8.14	4.43	123.2		$H^{\delta 1}$=6.58, $H^{\epsilon 1}$=10.37, $H^{\epsilon 3}$=7.28, $H^{\zeta 2}$=6.92, $H^{\zeta 3}$=5.70, $H^{\eta 2}$=5.39
V27	8.77	3.55	123.7	2.40	$H^{\gamma 1/\gamma 2}$=1.47/0.98
C28	7.68	2.02	120.2		
T29	8.34	3.58	117.8	4.12	$H^{\gamma 2}$=0.97
T30	8.69	4.12	113.4		
F31	7.77	3.80	125.2		$H^{\delta 1/\delta 2}$=5.91, $H^{\epsilon 1/\epsilon 2}$=7.17
H32	7.36	4.10	116.0	2.76	
T33	7.97				$H^{\gamma 2}$=1.40
S34					
G35	7.90	3.27/3.65			
Y36	7.12	3.90	106.1		
D37	7.08	5.14	118.7	2.49/3.15	
T38	8.49	4.07	115.4		
Q39	8.31	4.65	115.2	1.93/2.60	
A40	6.97	4.03	122.7	1.34	
I41	7.97	4.88		1.62	$H^{\gamma 12/\gamma 13/\gamma 2}$=0.67

Continued on next page

Continued from last page

Resid.	H^N	H^α	^{15}N	H^β/$H^{\beta 2/\beta 3}$	Other resonances
V42	8.36	4.25	127.9	2.15	$H^{\gamma 1/\gamma 2}$=0.87
Q43	8.68	4.85	126.7	2.07/1.93	$H^{\gamma 2/\gamma 3}$=2.33
N44	8.72	4.99	124.0	2.84/2.90	
N45	8.91	4.35	118.9	2.91	
D46	8.51	4.70	117.3	2.66/1.89	
S47	7.73	4.90	113.2	4.01	
T48	8.49	5.02	118.5	3.89	$H^{\gamma 2}$=0.64
E49	8.76	4.68	124.8		$H^{\gamma 2/\gamma 3}$=2.16
Y50	8.45	4.70			$H^{\delta 1/\delta 2}$=7.21, $H^{\epsilon 1/\epsilon 2}$=6.71
G51	9.38	4.32/4.97	112.4		
L52	10.14	3.94	123.8	1.25	H^{γ}=0.98, $H^{\delta 1/\delta 2}$=0.31/0.69
F53	8.75	4.54	109.3	1.92/3.37	$H^{\delta 1/\delta 2}$=6.91, $H^{\epsilon 1/\epsilon 2/\zeta}$=6.39/6.51,
Q54	8.05	3.68	117.1		$H^{\epsilon 21/\epsilon 22}$=5.35
I55	8.21	3.73	124.4	1.64	$H^{\gamma 12/\gamma 13}$=0.42/1.45, $H^{\gamma 2}$=1.06, $H^{\delta 1}$=-0.07
N56	7.48	5.93	125.8	3.20/3.37	
N57	8.42	5.34	118.1	3.56/3.59	$H^{\delta 21/\delta 22}$=10.00
K58	8.87	4.35	119.4		$H^{\beta 2/\beta 3/\gamma 2/\gamma 3/\delta 2/\delta 3/\epsilon 2/\epsilon 3}$=1.53/1.64/1.77/1.90/2.95
I59	7.47	4.29	111.2	0.79	$H^{\gamma 12/\gamma 13}$=0.96, $H^{\gamma 2}$=-0.16, $H^{\delta 1}$=0.62
W60	6.49	4.56	118.3	3.32/3.65	$H^{\delta 1}$=7.41, $H^{\epsilon 1}$=9.81, $H^{\epsilon 3}$=7.62, $H^{\zeta 2}$=7.27, $H^{\zeta 3}$=6.68, $H^{\eta 2}$=6.97
C61	7.91	5.82	111.8	3.07/2.89	
K62	8.19	4.94	122.6		
D63	8.33	4.77	127.8		
D64	8.13	4.45	114.9	2.60/2.65	
Q65	8.24	3.85	119.1	1.76	$H^{\gamma 2/\gamma 3}$=1.43
N66	8.07	5.37	114.3	2.67/3.19	
P67					$H^{\gamma 2/\gamma 3}$=3.53/3.47, $H^{\delta 2/\delta 3}$=2.01
H68	7.48	3.31			
S69	7.05	4.40	113.3	3.41/4.15	
S70	8.08	4.21	122.0	3.90/4.00	
N71	7.97	3.66	117.0	3.12	
I72	8.71	3.63	120.0	1.79	$H^{\gamma 12/\gamma 13}$=1.26, $H^{\gamma 2}$=1.05, $H^{\delta 1}$=0.43
C73	9.45	4.29	113.5	3.58/3.65	
N74	7.96	4.17	120.4	3.20	
I75	9.31	4.35	121.4		$H^{\gamma 12/\gamma 13}$=1.60, $H^{\gamma 2}$=0.76, $H^{\delta 1}$=0.43
S76	8.71	5.00	119.9	3.80	
C77	8.45		123.3		
D78	8.13		115.4		
K79	7.56	4.12	120.5		
F80	8.17	5.63	121.7	3.16	$H^{\delta 1/\delta 2}$=7.06, $H^{\epsilon 1/\epsilon 2}$=6.91
L81	7.00	5.73	113.7	1.86/1.93	H^{γ}=1.76
D82	7.44	4.87	118.6	2.74	

Continued on next page

Continued from last page

Resid.	H^N	H^α	^{15}N	H^β/$H^{\beta 2/\beta 3}$	Other resonances
D83	7.99	4.77		2.68/2.88	
D84	7.39	4.96	117.1	2.55/3.18	
L85	9.56	4.77	129.2	1.90	H^γ=1.72
T86	8.51	3.73	115.4	4.11	$H^{\gamma 2}$=1.28
D87	8.13	4.01	121.3	2.66/2.71	
D88	10.35	3.97		2.85	
I89	8.67	2.84	117.2	1.87	$H^{\gamma 12/\gamma 13}$=0.53, $H^{\delta 1}$=0.84
M90	7.52	3.90		2.27	$H^{\gamma 2/\gamma 3}$=2.62/2.81
C91	7.88	4.80	117.8	2.70/3.39	
V92	8.42	3.10	122.0	1.49	$H^{\gamma 1/\gamma 2}$=0.19/-0.49
K93	7.80	3.52	117.0	1.59	
K94	7.03	4.16	117.5		
I95	8.05	3.07	121.5	1.58	$H^{\gamma 12/\gamma 13}$=-2.41/0.82, $H^{\gamma 2}$=-0.16, $H^{\delta 1}$=-0.56
L96	9.08	3.83	121.7	2.28	H^γ=2.07, $H^{\delta 1/\delta 2}$=1.13/1.18
D97	8.40	4.46	118.1	2.79/2.99	
K98	7.96	4.53	117.9	1.84/2.23	$H^{\gamma 2/\gamma 3}$=1.74, $H^{\epsilon 2/\epsilon 3}$=3.05
V99	8.90	4.50	115.3	2.39	$H^{\gamma 1/\gamma 2}$=1.16/1.25
G100	8.20	4.29	109.9		
I101	8.30	4.42	117.6	1.49	$H^{\delta 1}$=0.21
N102	8.27	4.64	117.2	2.88	
Y103	7.60	3.86	120.3	2.12/2.59	$H^{\delta 1/\delta 2}$=5.50, $H^{\epsilon 1/\epsilon 2}$=6.55
W104	8.09	4.42	117.2		$H^{\delta 1}$=7.45, $H^{\epsilon 1}$=11.39, $H^{\epsilon 3}$=8.27, $H^{\zeta 2}$=7.04, $H^{\zeta 3}$=6.98, $H^{\eta 2}$=6.85
L105					
A106	8.75	4.06	116.6	1.37	
H107	7.82	4.83	115.2	3.16/3.46	
K108					
A109					
L110	7.62	4.61	113.8	1.60	
C111	7.32	5.13	116.6	2.73/3.27	
S112	8.08	4.55	111.3	3.74/3.86	
E113	7.71	4.59	118.3	1.97	$H^{\gamma 2/\gamma 3}$=2.20/2.12
K114	8.80	4.14	118.4		
L115	8.44	4.53	119.1		
D116	8.41	4.24	117.5	2.62/2.66	
Q117	7.91	4.06	115.9		
W118	7.77	4.58	121.2	3.09	$H^{\delta 1}$=7.27, $H^{\epsilon 1}$=10.70, $H^{\epsilon 3}$=7.41, $H^{\zeta 2}$=7.68, $H^{\zeta 3}$=6.79, $H^{\eta 2}$=6.52
L119	6.93	4.28	115.3	1.65	H^γ=1.87, $H^{\delta 1/\delta 2}$=1.09/1.55
C120	6.63	4.53	118.5	2.24/2.62	
E121	8.88	4.84	123.9	2.04/2.19	$H^{\gamma 2/\gamma 3}$=2.34/2.42
K122	7.92	4.33	121.0	1.86	$H^{\gamma 2/\gamma 3/\delta 2/\delta 3/\epsilon 2/\epsilon 3}$=1.41/1.73
L123	7.73	4.20	129.6	1.57	$H^{\gamma/\delta 1/\delta 2}$=0.88/0.91