5 Protein structure determination by NMR-spectroscopy

5.1 Structural studies by solution NMR spectroscopy

5.1.1 Preamble

Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique capable of determining high-resolution structures of biological macromolecules such as proteins and nucleic acids at atomic resolution. The magnetic resonance occurs as a result of the spin as quantum mechanical property. This is a source of angular momentum intrinsic to a number of different atomic nuclei. The spin angular momentum confers a magnetic moment on a nucleus and therefore a given energy in a magnetic field. The nuclear spin I must have a value different of zero (e.g., I=½) l, etc.) to give NMR spectra. The nuclei of greatest interest for biochemistry, H, C, N and P, all have isotopes with spin ½ having a magnetic moment, though for C and N the natural abundance of these isotopes (^{13}C, ^{15}N) is very small. However, these isotopes can be incorporated into biomolecules through isotope labelling.
In the following, the basic principles of NMR are described and the applications of biomolecular NMR in structural biology are introduced.

5.1.2 Theory of NMR spectroscopy

When a sample (nuclei with spin ½) is placed in a magnetic field B_0, the macroscopic magnetization lies parallel to B_0 (i. e. along the z-axis of the rotating frame). To record a conventional 1D NMR spectrum, a radiofrequency pulse B_1 is applied that rotates the magnetization away from the z-axis toward the x,y plane. The free induction decay (FID) is then recorded immediately after the pulse. Fourier transformation (FT) of the FID yields the conventional 1D spectrum. To obtain additional information on interactions between spins double or multiple irradiation experiments must be carried out.
In the microscopic view, the nuclei with spin ½ are aligned either in a parallel or antiparallel orientation with respect to the magnetic field. The parallel α state is lower in energy than the antiparallel β state. The energy of these levels is given by the classical formula for a magnetic dipole in a homogeneous magnetic field of the strength B_0:

$$E= -\mu_z * B_0= -m * \gamma * h/(2\pi) * B_0 \quad (5.1)$$

where h, the Planck's constant; γ, the gyromagnetic ratio; m, the magnetic quantum number, an integer number between –I and +I.

Therefore, the population of the α state is greater than that of the β state (Boltzmann distribution), and a macroscopic magnetization along the B_0 field exists due to this population difference. The magnetic moment of each nucleus precesses around B_0. The frequency of this precession is the larmor frequency (ω_0) which is equivalent to the energy difference between the two levels. In order to measure the precession frequencies, the system is disturbed and brought into a non-equilibrium state which allows monitoring the transverse (x or y) magnetization. This can be done by applying an oscillatory magnetic field B1 that rotates the magnetization away from the z-axis toward the x, y plane. Being composed of individual nuclear spins, a transverse (in the x, y plane) macroscopic magnetization $M_{x,y}$ (coherence) starts precessing about the z axis with the Larmor frequency (in the lab coordinate system) under the influence of the static B_0 field, thus inducing a current in the receiver coil, which is then recorded and called FID. The FID obtained from a FT NMR experiment is a superposition of the frequencies of all spins in the molecule as a function of time, F(t). In order to obtain the corresponding spectrum F(ω) (intensity as a function of frequency) a Fourier transformation is performed, which is a mathematical operation which translates a function in the time domain into the frequency domain.

Relaxation and NMR line widths. After a 90° pulse has been applied to the equilibrium z-magnetization, a transverse magnetization is created in which the phases of the individual spin precession frequencies are correlated. Such a non-random superposition of states is called coherence. However, the population difference between the α- and β-states (which corresponds to the z-magnetization) has disappeared. Thus, the system is in a non-equilibrium state and will therefore relax back into the equilibrium state. This is described by two relaxation processes. The loss of phase-coherence in the transverse plane due to spin-spin interactions defines the T2 or transverse relaxation time. In addition, the population difference along z is restored by interactions with the surroundings. This spin-lattice relaxation time is described by the exponential decay of the FID, and reciprocal to the transverse relaxation time T2. It scales with the molecular tumbling rate in solution and therefore increases with higher molecular weight. This is the main reason for the molecular weight limitation of NMR.

Chemical shifts. The static magnetic field B_0 is shielded by the electrons in the local environment of a spin. Therefore, the individual resonance frequencies are slighty different reflecting the different chemical surroundings. The resonance frequencies are called chemical shifts and are measured in parts per million (ppm) in order to have chemical shift values independent of the static magnetic field strength. For example, backbone amide proton HN in a proton resonate around 8 ppm, while Hα spins have resonance frequencies between 3.5-5.5 ppm.

J-coupling. Scalar or J-couplings are mediated through chemical bonds connecting two spins. The energy levels of each spin are slightly altered depending on the spin state of a scalar coupled spin (α or β). This gives rise to a splitting of the resonance lines. Typical coupling constants are: ^{1}J (HN, N) ~92 Hz, ^{3}J(HN, Hα) ~ 2-10 Hz, $^{>3}$J < 1 Hz. ^{3}J-couplings are well-correlated with the central dihedral angle by an empirical correlation, the Karplus curve. For example, ^{3}J(HN, Hα) defines the backbone angle φ in proteins. Scalar couplings are used in multidimensiolnal (2D, 3D, 4D) correlation experiments to transfer magnetization from one than three chemical bonds. For example, the amino acid ^{1}H spin system of alanine consists of HN, Hα and Hβ.

Nuclear Overhauser effect (NOE). The nuclear Overhauser effect is a result of cross relaxation between dipolar coupled spins as a result of spin/spin interactions through space. Dipolar couplings are usually in the kHz range, and depend on the distance between the two spins and the orientation of the internuclear vector with respect to the static magnetic field B_0. Due to the fast overall tumbling of molecules in solution, the dipolar couplings are averaged to zero. Nevertheless, the dipolar couplings give rise to spin/spin and spin/lattice relaxation. The NOE is a result of cross-relaxation between spins and is defined by the transition rates W_0 and W_2 which involve spin flips of both spins. The NOE allows to transfer magnetization from one spin to another through space, and scales with the

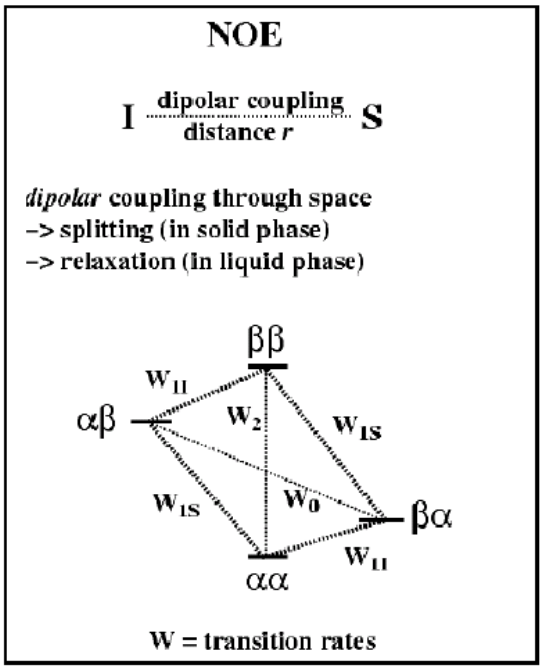

Fig. 5.1 Illustration of NOE.

distance r between the two spins (NOE ~ $1/r^6$), e. g. two protons in a protein. Therefore the NOE is related to the 3D-structure of a molecule. For interproton distances >5 Å, the NOE is too small and not observable.

Residual dipolar couplings (RDC). For two directly coupled nuclei A and B, the direct dipolar coupling interaction is quantified by

$$D = \frac{-\mu_0 \gamma_A \gamma_B h}{8\pi} \; r^{-3}_{AB} \; \frac{3\cos^2\theta - 1}{2} \qquad (5.2)$$

where μ_0 is the permeability of free space, γ_A and γ_B are the gyromagnetic ratios of the coupled nuclei, and $\langle r^{-3}_{AB} \rangle$ is the motionally-averaged inverse cube of the internuclear distance, θ is the time-dependent angle between a unit vector in the internuclear direction and a unit vector parallel to B_0, and the brackets signify the time average of the quantity. Normally, the random isotropic sampling of angles by a molecule tumbling in solution reduces the RDC to zero. This isotropic sampling may be made anisotropic by a magnetically induced alignment or with the aid of various types of media. This anisotropic sampling will result in a measurable RDC that is indicative of the average orientation of an inter-nuclear vector.

Limitation by signal overlapping, can often be overcome by extending the measurements into a second dimension. All 2D NMR experiments comprise the same basic scheme: exciting the nuclei with two pulses or group of pulses then receiving the FID. The acquisition is carried out many times, incrementing the delay (evolution time t_1) between the two pulse groups. Any 2D NMR experiments consist of a preparation period, an evolution period (t_1) during which the spins are labelled according to their chemical shift, a mixing period during which the spins are correlated with each other, and finally a detection period (t_2). 2D FT of the [t_1, t_2] data matrix yields the desired 2D frequency spectrum. A pulse sequence for a typical 2D experiment is shown (Fig. 5.2 A).

The nature of the interaction during the mixing time depends on the type of experiment. Thus, in correlated spectroscopy (COSY) and total correlated spectroscopy (TOCSY) experiments the cross-peaks arise from through-bond scalar correlations (Fig. 5.2 C), while in a nuclear Overhauser enhancement spectroscopy (NOESY) experiment (Fig. 5.2 B) they arise from through-space correlations.

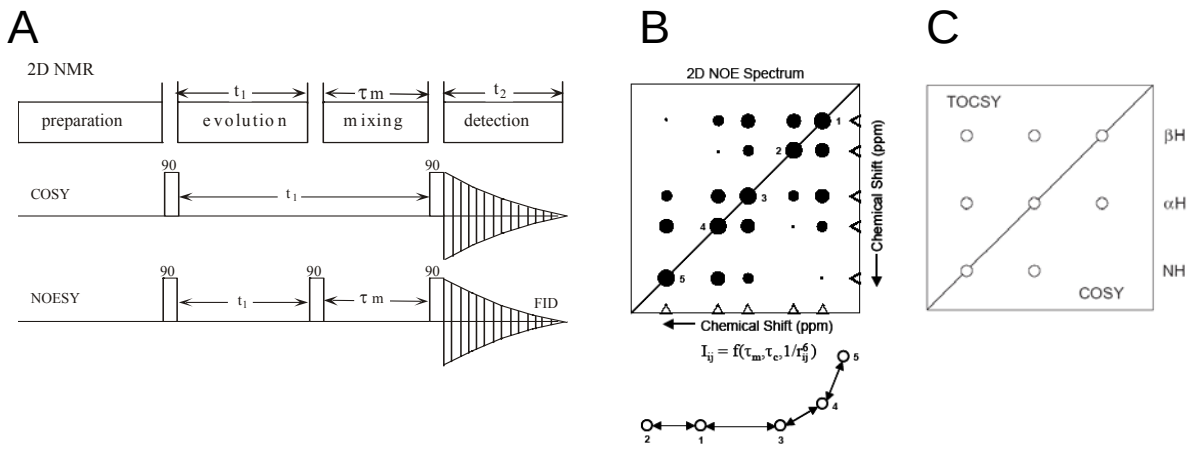

Fig 5.2 (A) The pulse sequence for 2D NOESY and COSY experiments. (B) hypothetic 2D NOESY spectrum. The intensities of the observed out-of-diagonal cross correlations between two protons are a function of the rate of the molecular tumbling and of the inverse sixth power of the distance separating those protons. (C) Simulated TOCSY and COSY spectra of an alanine spin system with three spins NH, αH and βH.

The aim of COSY experiments is to correlate signals of different nuclei (generally protons separated by two or three bonds) by a scalar coupling through bonds. ^{1}H-^{1}H TOCSY is useful for determining which signals arise from protons within a spin system. During the mixing time the magnetization is spin-locked, which results in coherence transfer between all coupled nuclei in a spin system (even if they are not directly coupled). In contrast, NOEs result from cross-relaxation effects between protons, observed only for pairs of protons that are separated by less than 5 Å, regardless of covalent structure.

Two features of these spectra are clearly evident. The first is the extensive overlap of resonances that renders unambiguous interpretation of cross-peaks virtually impossible for larger proteins. The second is that the number of connectivities in the TOCSY spectrum form the HN protons to the Hβ and beyond is small. This is due to the fact that the homonuclear J-couplings that form the basis of correlation experiments (e. g., ^{1}H-^{1}H COSY, TOCSY, etc.) are small. Moreover, these couplings are often unresolved because of the relatively large linewidths of protein resonances, which are an inevitable consequence of the slower rotational correlation time of the protein as the molecular weight increases. As a result, assignment strategies based on conventional 2D NMR experiments break down for proteins the size of >90 amino acids.

The assignment strategy proposed by Wüthrich in 1986 concerns the molecules in natural abundance. It is based on the proton study via 2D experiments such as COSY, TOCSY, NOESY and ROESY to solve relative small protein structures of less that about 10 kDa. The proton spin system of each amino acid residue is identified from the COSY and TOCSY

spectra and the information provided by the NOESY/ROESY spectra is used for the sequential assignment and converted into distance restraints.
Problems associated with chemical shift overlap in 2D NMR spectra can sometimes be overcome by increasing the dimensionality of the spectrum to 3D.

5.1.3 Structure determination of [U-^{15}N, ^{13}C]labelled proteins

5.1.3.1 Introduction

The problems of limited resolution among the Hα resonances and the conformation dependence of the NOE and ^{3}J coupling constants are overcome in the assignment strategies for ^{15}N, ^{13}C-labelled proteins, since they employ coherence transfer via ^{1}J (and sometimes ^{2}J) couplings only, which are largely independent of conformation. The general approach for solving the solution structure with double labelled proteins can be summarized as follows: preparation of protein solution, NMR measurements, assignment of NMR signals to individual atoms in the molecule, identification of conformational constraints (e. g. distances between H-atoms), the calculation of the 3D structure on the basis of the experimental constraints. To succeed in structure determination of double labelled proteins, the chemical shift assignment must be nearly completely achieved, which is the crucial point. In this way, 3D NMR permits to use a large number of triple resonance experiments to assign the backbone and side-chain resonances. The nomenclature for these experiments reflects the magnetization transfer pathway. The spins, whose chemical shifts are not evolved, are put in parentheses. In the following section, NMR experiments for assignment of ^{15}N, ^{13}C-labelled proteins are summarized.

5.1.3.2 Assignment of the backbone chemical shifts

A strategy based on triple resonance experiments employs several 3D experiments to correlate the resonances of the peptide backbone (HN(i), NH(i), Cα(i), Hα(i), Cα(i-1), Hα(i-1), C'(i) and C'(i-1)).
Almost all experiments contain the ^{15}N- and HN-resonances, therefore allowing the use of this pair of spins as reference and starting point for further assignment of other resonances. The HNCA experiment, for example, correlates the HN and ^{15}N chemical shifts of residue (i) with ^{13}Cα shifts of residue (i) (via $^{1}J_{NC\alpha}$= 7-12 Hz) and residue (i-1) (via $^{2}J_{NC\alpha}$< 8 Hz), thereby providing sequential connectivity information. A complementary experiment to the HNCA, the HN(CO)CA correlates, in contrast, the amide proton and nitrogen resonances of an amino acid only with the Cα chemical shift of its preceding residue. This is due to the fact that this

technique uses a relay mechanism, transferring magnetization from ^{15}N to ^{13}C via the intermediate carbonyl nucleus. Generally, the intensity of the Cα(i) and Cα(i-1) of the HNCA cross peaks can be differentiated on the basis of their relative intensity. Thus the HN(CO)CA proposes only an unambiguous assignment in case of accidental overlap of intra- and interresidue HN-N- Cα correlations in the HNCA experiment. An HNHA experiment completes the assignment of the α-resonances. The spins Hα(i) are correlated to the HN via a $^{3}J_{H}N_{H\alpha}(i)$ coupling that connects only the resonances of the same residue.

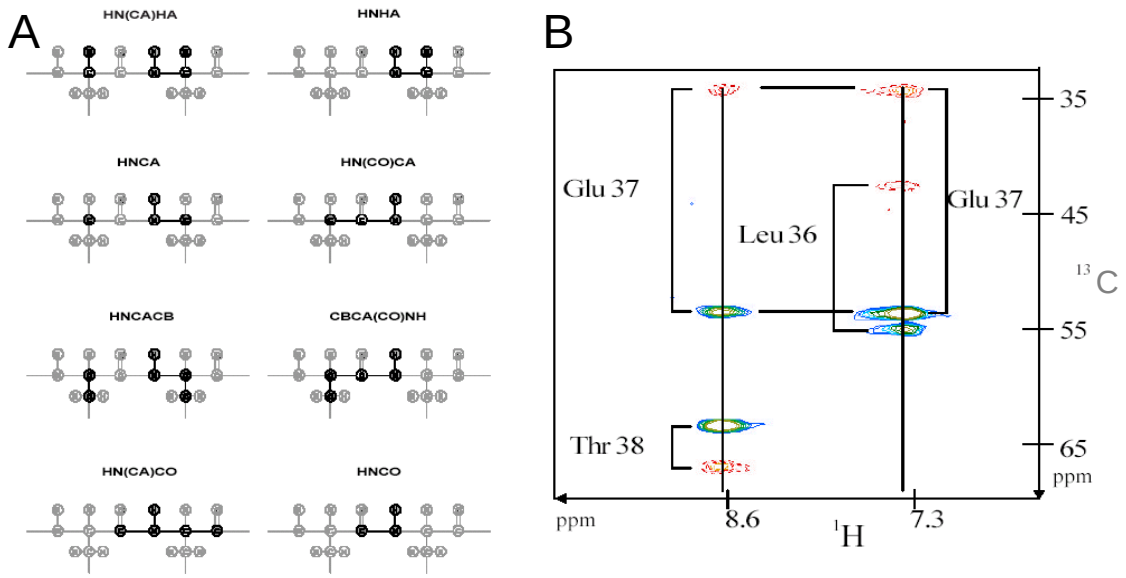

Fig. 5.3 (A) Combination of triple resonance experiments for the backbone assignment of doubly ^{15}N-^{13}C labelled proteins. (B) Two spin system traces showing the four aliphatic carbon correlations in a HNCACB. The Cα(i) and Cα(i-1) are coloured in blue; the Cβ(i) and Cβ(i-1) are coloured in red.

To circumvent the severe overlap of the α-resonances, further developments involved the chemical shifts of side-chain carbon and proton spins (especially Cβ and Hβ) to achieve the sequential assignment. It follows then the recommended set of experiments (Fig. 5.3 A): HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HNHA, HN(CA)HA, HNHB. Among these atoms, Cα and Cβ are of prime importance, first because their chemical shifts show a large spectral dispersion (Cα ≈ 25 ppm; Cβ ≈ 60 ppm), and then because taken together, these shifts are characteristic for the identification of the amino acids.

The most significant experiment for the assignment of the backbone resonances is the HNCACB. This experiment yields the Cβ shifts (in position (i) and (i-1)) in addition to those coming from the HNCA. The Cα and Cβ correlations have opposite signs and can thus be distinguished (Fig. 5.3 B). The resonances in the (i)-position can be discriminated from those in the (i-1)-position on the basis of their different intensity as explained for the HNCA experiment.

The HNCACB allows a complete assignment of the Cα and Cβ resonances experiment, providing together the recognition of the residue and sequential information. To complete the β-resonances assignment, and HNHB gives exclusively the Hβ shifts via a $^{3}J_{NH\beta}$ coupling, and similarly to the pair HNCA/HN(CO)CA, the CBCA(CO)NH provides specifically the sequential information, i. e. the Cα/Cβ (i-1) only.

5.1.4 Structural studies by solid-state NMR spectroscopy – An introduction

Over the past two decades, solid-state NMR has emerged as a powerful method for structural studies of solid biomolecules, such as membrane proteins and amyloid systems. In contrast to that in solution, fast molecular tumbling in most solids is absent and signals in the spectrum is extremely broad (≈ 150 ppm) and generally uninterpretable due to three major anisotropic interactions: chemical shift anisotropy (CSA), homonuclear dipolar interactions and heteronuclear dipolar couplings of spin ½ nuclei. To overcome this problem, the magic-angle spinning (MAS) technique has been developed to obtain high resolution spectra, where the spectral resolution is highly improved by mechanically rotating the sample at the magic angle, that is around an axis tilted of 54.7° relative to the external magnetic field. The rotation about this axis removes the broadening and leads to a considerable sharpening of the lines in spectrum. A review can be found in Laws *et al*. (2002). MAS introduced NMR spectra provides sufficiently resolved peaks and allows to extract the information about the structure or dynamics of the molecule. This has important advantages for structural studies by solid-state MAS NMR. First, well established strategies for resonance assignment known from solution NMR can be implemented in solid-state MAS NMR, and chemical shift databases compiled from solution NMR studies can be accessed to identify amino-acids by means of their characteristic side-chain correlation patterns.

5.1.4.1 Basic solid-state NMR techniques

Magic angle spinning

One approach to eliminate the interactions to give narrower lines is magic angle sample spinning. This is done because the value of $(3\cos^2\theta - 1)$ term is zero when the vector between two nuclei makes an angle θ=54.74° (the magic angle) with the static magnetic field as shown in Fig. 5.4A. Therefore, chemical shift anisotropy, dipolar couplings and quadrupolar interactions can be averaged to zero. J couplings is on the order of 10^2 Hz, is much smaller to be observable in solid-state NMR.

The effect of MAS on different Hamiltonians, homogeneous and inhomogeneous Hamiltonians are introduced here. Anisotropic interactions can be divided into inhomogeneous interactions including the chemical shift, quadrupolar interaction and heteronuclear dipolar coupling and homogeneous one referring to homonuclear dipolar interactions. In the inhomogeneous case, e. g., in a ^{13}C spectrum, where the CSA interaction overrides the homonuclear dipolar interaction, the broad line splits up into many narrow lines which are spaced by the spinning frequency under MAS at a low spinning frequency. If the spinning speed (ω_r) is increased and exceeds the magnitude of interaction, the interactions will be completely averaged and the spacing of the lines gets larger and finally end up with a single line at the isotropic chemical shift (Fig. 5.4B). The same effect of MAS is presented on a spectrum where the heteronuclear dipolar coupling interaction is the dominant interaction experienced by the spin. If the spinning speed fails to completely average the anisotropic interactions, the remaining anisotropic contribution leads to sidebands spaced at intervals, equalling the rotor period (ω_r). In order to avoid reduction in sensitivity, the sidebands can be minimized by spinning at a frequency greater than the width of the CSA pattern. For the homogeneous case, the same pattern of spinning sidebands could be obtained for a single homonuclear dipolar coupling interaction. However, for a network of multiple homonuclear dipolar couplings, the pattern changes.

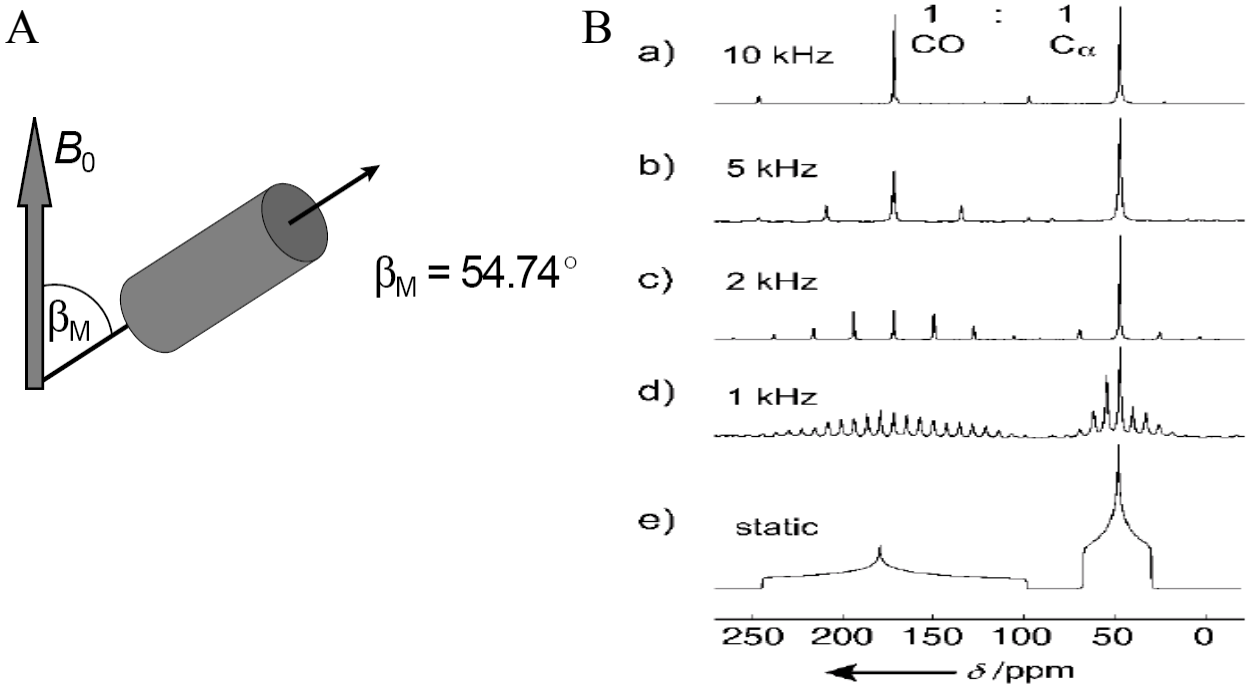

Fig. 5.4 (A) By rotating about the magic angle the time-average value of all binding vectors become 54°74' relative to the main field. The CSA is minimized. (B) Solid-state ^{13}C NMR spectra (125 MHz) of a uniformly ^{13}C-labelled (10%) glycine powder showing the effect of MAS on the anisotropic lineshape due to CSA interaction. The spectra (a-d) were acquired under CW ^{1}H decoupling and MAS at the given spinning speeds. Spectrum e represents a powder spectrum reconstructed from the CSA parameters obtained by fitting the sideband intensities in spectra a-d.

Cross Polarization

The direct excitation and detection of spins such as ^{13}C and ^{15}N with a low gyromagnetic ratio is plagued by low sensitivity. To enhance the signals from rare nuclei, many solid-state NMR experiments routinely involve the transfer of polarization from abundant nuclei (usually ^{1}H nuclei) by using a technique called cross polarization (CP) originally introduced by Hartmann and Hahn. The process of CP occurs through the tendency of the magnetization to flow from highly polarized nuclei to nuclei with lower polarizations when two nuclei are brought into contact. The exchange of magnetization must be driven externally by the application of RF fields. The Hartmann-Hahn method requires the simultaneous application of two continuous RF fields, one at the resonance frequency of the spin I and one at the resonance frequency of spin S. The effect of any RF field is to rotate the magnetization about the axis of the applied field. The rotation rate depends on the frequency and amplitude of the RF field. An RF field that oscillates at the spin I frequency, e. g. 500 MHz, would have essentially no effect on spin S with a frequency of 125 MHz and *vice versa*. By applying two RF fields, one tuned to the spin I and the other to the spin S, both the I and S spins can be rotated independently around a particular axis at rates determined by the amplitudes of the two applied fields. When the rotation frequencies of the spins I and S are equal, an energy conserving dipolar contact between the two spin systems is created. The differences in energy are supplied by the RF fields. One way to describe how this connection is established is by using a reference frame rotating at both the I and S spin rotation frequencies. In the double rotating frame, the spacing between the spin-up and spin-down energy levels is equal for the spins I and S.

The experimental implementation of this concept to obtain high-resolution NMR spectra of rare or dilute S spins is shown in Fig. 5.5. First, the proton magnetization is brought into the xy plane by a non selective $\pi/2$ pulse. RF fields are then applied to the I and S spins for a period τ_{CP}, causing the magnetization to be exchanged between the spins I and S in their respective rotating frame. Finally, the spins S are detected while spins I are decoupled. The increase in the spin S magnetization during the CP mixing period τ_{CP} depends on the strength of the I-S dipolar coupling. The polarization buildup in the unprotonated carbonyl carbon atom requires a longer CP time than for the α carbon atom, which is quickly polarized by its bonded protons.

The magnitude of this exchange depends on the magnitude of B_I and B_S and is maximal when the Hartmann-Hahn condition is fulfilled: $\omega_I = \omega_S$ ($\gamma_I B_I = \gamma_S B_S$). ($\gamma_I$ and γ_S are gyromagnetic ratios of spins I and S, respectively; B_I and B_S are nutation rates imposed by the RF fields). The methodology offers important gains, as the sensitivity of the observed nucleus is

increased according to the ratio of the gyromagnetic ratios, offering a 4 and 10 fold theoretical improvement in sensitivity for the observation of ^{13}C and ^{15}N, respectively.

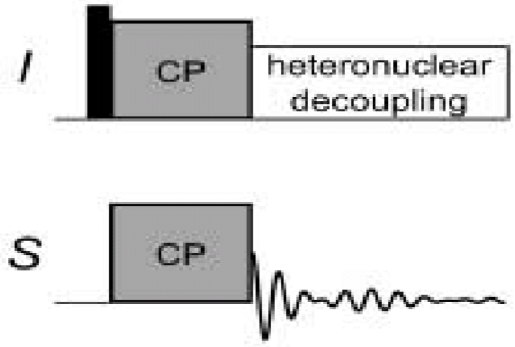

Fig. 5.5 Pulse sequence for CP of ^{13}C (S) nuclei by proton (I) with detection of the ^{13}C magnetization.

Dipolar recoupling

The direct dipolar couplings are very informative regarding the structure of a molecule. It is not possible to directly obtain the dipolar couplings from a normal CP/MAS experiment, as this experiment is designed to remove the dipolar couplings together with CSA and part of the quadrupolar interactions. It is however possible to reintroduce the dipolar couplings, both the hetero- and the homonuclear couplings, by using different pulse sequences, often synchronisized with the sample rotation.

Homonuclear dipolar couplings

The detection of homonuclear dipolar couplings in solid state NMR was first accomplished by Griffin and co-workers. They discovered that when the spinning speed is adjusted so that a spinning sideband from one nucleus and the signal from another in its vicinity coincide, the intensity of that signal changes. This phenomenon is called rotational resonance and is one of two fundamentally different methods to measure homonuclear dipolar couplings. This method suffers from a restriction of the nuclei that are to be studied. If the chemical shift difference is too small, it is not possible to use MAS speeds to match the difference without other interactions, e.g. CSA, becoming significant and leading to broad signals that will be difficult to interpret. If the chemical shift difference is too large, i. e. larger than the magnitude of the dipolar coupling in Hz, the method again fails, as there will be no SS connecting the signals from the coupled nuclei. If the chemical shift difference is small, a different approach using double quantum coherence (DQC) is more suitable. This was introduced by Tycko and co-workers, who showed that rf pulses applied synchroized with the rotor period have the effect of interrupting the dipolar averaging.

Heteronuclear dipolar couplings

Heteronuclear dipolar couplings between isolated spins can be studied by the Rotational Echoe Double Resonance, REDOR experiment, introduced by Gullion & Schaefer (1989). This method uses rotor-synchronized π-pulses to prevent the averaging of the heteronuclear couplings in much the same way as in the homonuclear experiments described above. In the REDOR experiment (Fig. 5.6) S spin magnetization is first prepared, followed by an I-S dipolar evolution period and finally the S spin signal is detected. The dipolar evolution period is built from a series of π-pulses on the I spin, and one S spin π-pulse applied at the centre of the evolution period. This S spin π-pulse is applied to ensure that the S spin magnetization is refocused at the beginning of the acquisition. By measuring the REDOR signal decay with and without the π-pulses, the dipolar coupling constant can be derived. This coupling constant is strongly dependent on the internuclear distance between the two coupled nuclei and the distance can be determined quite accurately.

$$D = \frac{\mu_0 \gamma_I \gamma_S h}{4\pi r^3_{IS}} \qquad (5.3)$$

where γ_x is the gyromagnetic ratio for spin I and S, r_{IS} is the interatomic distance between two spins and the other symbol have their usual meaning.

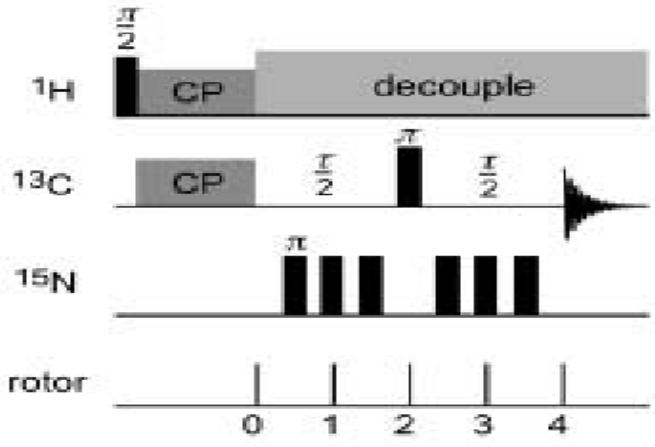

Fig. 5.6 The REDOR pulse sequence. CP of the ^{13}C nuclei (S spin) is followed by a period τ in which the ^{13}C magnetization is dephased by the application of a series of rotor-synchronized pulses on the ^{15}N (I spin) magnetization, which recouple the ^{13}C-^{15}N dipolar coupling.

5.1.5 Structure determination by solid-state MAS NMR

5.1.5.1 Solid state NMR for membrane proteins

The physical state of the membrane in which a protein is embedded is intermediate between that of an isotropic liquid and a solid. Natural membranes are liquid crystalline, typically exhibiting order in one dimension but disorder in the other two dimensions. Because of the anisotropic nature of the system, the orientationally dependent interactions of an NMR nucleus with its environment are not totally averaged by molecular reorientation (as occurs in isotropic liquids studied in solution state NMR).

Membrane proteins are uniquely oriented and have a high degree of anisotropy. However, if studied as a purified protein, together with a requisite detergent micelle, then motionally broadened spectra of insufficient resolution for structural analyses are recorded using conventional solution state NMR approaches.

Using solid state NMR techniques, there are two approaches that can be used to deal with the anisotropic interactions which lead to NMR spectral broadening and permit molecular structural information to be extracted. One is to exploit the spectral anisotropy in oriented samples to give molecular orientations in static samples, and the second is to use random dispersions in MAS NMR, in which orientational information may be lost, but can be regained from analysis of spinning sidebands. In addition, dipolar couplings can be recoupled to yield distance constraints, as well as chemical shift information and quadrupolar interactions to define local environmental characteristics.

Solid-state NMR allows a direct examination of a protein in the membrane, either as a single species (reconstituted in defined lipids) or in a heterogeneous environment with other proteins and lipids (natural membranes). A detailed description of orientational constraints with respect to the membrane can be obtained, including identification of secondary structural elements. The specific parts of a protein that are vital to ligand and drugs binding can be identified, due to chemical shift data which can define pharmacophores. To date, however, only a few membrane proteins or peptides have been studied whilst bound using solid-state NMR.

5.1.5.2 The methodology

The methodology has been previously developed in the group and consists of broad-band recoupling method and spin dilution. The small well-ordered microcrystalline SH3 domain of α-spectrin (62 residues) with known 3D structure and chemical shifts available from solution-state NMR have been used to develop and validate this solid-state NMR methodology. In a

first stage, 2D correlation experiments aiming at the spectral assignment of protein resonances in the solid state were performed. After obtaining nearly complete ^{13}C and ^{15}N sequential resonance assignments in the SH3 domain of α-spectrin (Pauli *et al*., 2001), the resulting resonance assignments provided the basis for the interpretation of nontrivial (^{15}N,^{15}N) and (^{13}C,^{13}C) contacts of a ^{13}C block labelled variant of the SH3 domain of α-spectrin and subsequent 3D structure calculation (Castellani *et al*., 2002).

The methodology is based on the detection of a large number of ^{13}C-^{13}C structure defining distances from a small number of samples. For side chain correlation and assignment experiments, it is desirable to achieve broad-banded (i. e. non-spectrally selective) efficient recoupling. In a uniformly labelled preparation, the detection of long-range ^{13}C-^{13}C restraints is obscured by the fact that the strong couplings between directly bonded spins (e. g. ^{13}C-^{13}C scalar couplings) would interfere with the weak couplings needed to get distance information in the range 3-6 Å. The effect is referred to dipolar truncation. With broadband dipolar recoupling techniques it is possible to perform spectral assignments in uniformly labelled molecules and to measure distances in molecules labelled with spin pairs. However, it is not possible to measure multiple distances, and therefore to constrain structures, in these systems with these techniques. The problem of detecting structurally relevant distance constraints in solid state protein can alternatively be solved by biochemistry, i. e., reduced isotope-labelling schemes, in which chemically-bonded carbons are mostly not simultaneously labelled. Dilution of the ^{13}C labelling level avoids the strong couplings and therefore broadband recoupling methods can be used. Then it is possible to detect weak couplings that contain structural information. A reduced ^{13}C-block labelling (refer to as spin dilution) can be achieved by using [1,3-^{13}C]- glycerol and [2-^{13}C]-glycerol as the sole nutrient carbon source in bacteria medium. The result is an almost alternating labelling scheme, where still many carbon sites are labelled but not in adjacent positions. In combination with this labelling pattern, long-range ^{13}C-^{13}C distance restraints are collected by using a broad-band recoupling method like the proton-driven spinning diffusion (PDSD) mixing scheme. It is anticipated that this strategy is applicable for study of receptor-bound agonists and antagonists. The disadvantage is that the signal/noise in the cross peaks is reduced over that observed in a uniformly labelled sample. Spectrally selective recoupling avoids the problem of dipolar truncation by recoupling the ^{13}C-^{13}C's one pair at a time.

5.1.5.3 Selective ^{13}C labelling

The labelling method relies on the specificity of the amino acid biosynthetic pathways in bacteria. Bacteria can utilize glucose or glycerol as the sole carbon source to synthesize the 20

amino acids. This is achieved via three major pathways: glycolysis, pentose phosphate pathway and citric acid cycle. In Fig. 5.7 the two groups of amino acids are shown, together with the effective ^{13}C labelling grade, as obtained by protein expression with [2-^{13}C] glycerol or [1,3-^{13}C] glycerol. The labelling scheme of group I is characterized by sites either 100% or 0% enriched. The opposite labelling is obtained for [1,3-^{13}C] glycerol. The amino acids of group II show a more complex labelling scheme.

Knowledge of the labelling patterns and labelling levels affords additional information on resonance assignment that would be difficult to extract from the spectra of fully ^{13}C labelled proteins. The combined use of signal intensities based on the known ^{13}C labelling levels and the characteristic chemical shift ranges can reduce the ambiguity of the amino acid type assignment.

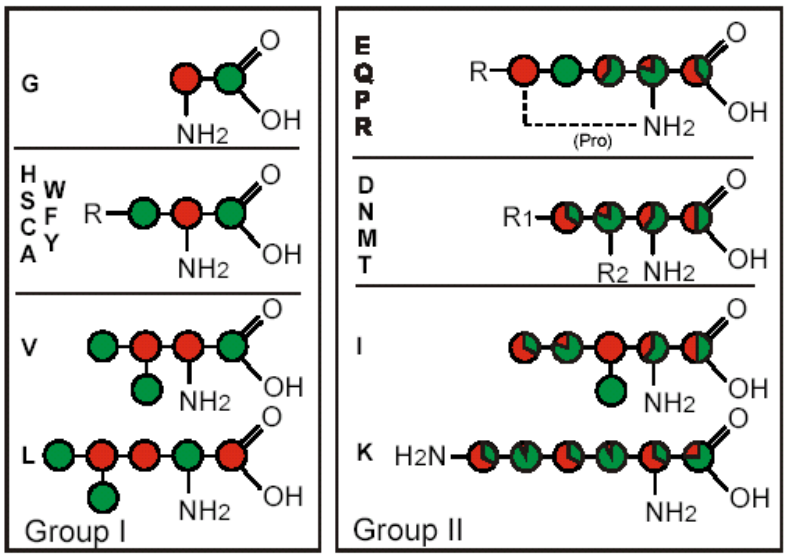

Fig. 5.7 Schematic representation of the effective ^{13}C enrichment, as obtained by protein expression in *E. coli*. The red colour corresponds to the degree of ^{13}C labelling obtained by growth on [2-^{13}C] glycerol. The opposite labelling pattern obtained by [1,3-^{13}C]glycerol is presented in green.

5.1.5.4 Assignment strategy and methods

The assignment strategy is based mainly on ^{13}C-^{13}C correlations, applied to several samples. More specifically, intra-residue correlations are obtained from a PDSD experiment with short mixing time and inter-residue correlations are obtained from a PDSD experiment with a long mixing time, which reveal ^{13}C connectivities as off-diagonal intensities. A detailed discussion of this pulse scheme is, e. g. given by Pauli *et al*. (2001). Intra-residue resonance assignment usually begins with homonuclear ^{13}C-^{13}C correlation experiments that allow identification of

different amino acid types based on their characteristic chemical shift correlation patterns. In the solid state, one can use not only scalar couplings but also dipolar couplings for coherence transfer. Due to its sharp distance dependence ($1/r^3$), the dipolar interaction can distinguish one-bond from multiple-bond connectivities quite effectively. Once the resonances in the NMR spectra have been assigned, structural constraints can be extracted to gain insight into the secondary structure of the peptide. To this end, three main classes of experiments can be used: experiments that measure conformation-dependent chemical shifts, experiments that yield torsion angles and, finally experiments where interatomic distances can be determined. Here, secondary structure information should be obtained from the semi-quantitative measurement of ^{13}C-^{13}C and ^{15}N-^{15}N distances, as described by Castellani *et al*. (2002).

5.1.6 Aim of study

Two general NMR strategies are available for structure determination. The first involves distance measurements on site-specifically labelled samples. The second approach uses more qualitative methods such as spin diffusion to determine the proximity of spins between different molecules. In this work, two receptor samples were available. The ligand-receptor sample for the REDOR experiment incorporated specific ^{13}C, ^{15}N labelled ET-1. The complex was prepared as described in Chapter 3. The goal here was to explore the feasibility of distance measurements if it can set the stage for distance measurements elsewhere in the receptor. The second ligand-receptor sample contained uniformly labelled ET-1. The primary goal of this study was to examine the feasibility of the above mentioned current 2D solid-state NMR approach for structural studies of ligands bound to receptors. Moreover, although isotopically labelled ET-1 for NMR studies of ligand-receptor interaction was synthesized previously (Chapter 4), there is no study yet on its actual application. To demonstrate the potential of the labelled peptide in NMR studies of ligand-receptor interaction, a 1:1 complex of the uniformly ^{13}C-labelled ET-1 with the ET_BR from Sf9 cells was prepared. Sufficient quantities of the receptor (milligram order) have become available, because of recent development of fermentation culturing techniques, and was provided.

5.2 Materials and experiments

5.2.1 ET-1 synthesis

The quantities of isotopically labelled ET-1 required for NMR spectroscopy were prepared by expression of recombinant peptides in *E. coli*, for partial or uniform isotopic labelling (Chapter 4), or by solid-phase synthesis, for specific residue labelling of the samples. The synthesis and purification of synthetically labelled ET-1 was performed at the FMP Berlin in the group of Dr. M. Beyermann . ET-1 was synthesized by solid phase peptide synthesis using Fmoc (9-fluorenylmethoxycarbonyl) chemistry.

5.2.2 Receptor preparation

The details of the receptor preparation from *P. pastoris* for REDOR experiments are given in Chapter 3. The ET_BR protein prepared from Sf9 cells was expressed by an industrial company using a rocking wave fermenter. Detergent solubilized ET_BR from Sf9 cells was obtained from A. Srivastava (MPI for Biophysics, Frankfurt a. Main), and its preparation details are not given. The recombinant ET_BR was purified from a portion of a 100 l Sf9 fermenter culture (~14 mg purified ET_BR/100 l; A. Srivastava, personal communication). The ligand-free receptor after purification was in detergent buffer composed of 20 mM Hepes, pH 7.4, 150 mM NaCl, 0.01% n-dodecyl-maltoside. Prior to NMR measurement a stoichiometric quantity of U-^{13}C,^{15}N-ET-1 was added to the receptor sample. Upon the addition of a stoichiometric amount of ET-1 the complex was formed and excess of free ET-1 was removed by washing. The receptor complex was concentrated using a Centricon containing a 50 kDa cut off membrane. The NMR sample consisted of about 80 nmol (4 mg) receptor.

5.2.3 Solid-state NMR

1D MAS ^{13}C NMR spectra of ET_BR/ET-1 purified from *P. pastoris* were obtained with a Bruker DMX-600 spectrometer. A 4-mm triple resonance MAS probe was used together with a rotor containing an inner spacer in order to improve B_0 homogeneity and spinning stability. 1D ^{13}C spectra were acquired at 220 K and at MAS frequencies of 9 kHz. Magnetization was transferred from ^{1}H to ^{13}C with a ramped CP contact of 1 ms. 2D PDSD MAS ^{13}C-^{13}C NMR spectra were recorded on a Bruker 600 MHz spectrometer. MAS frequency was set to 9.5 kHz to minimize sideband overlap and to avoid rotational resonance effects between directly

bonded ^{13}C labels. TPPM of 85 kHz was used for heteronuclear decoupling in both dimensions. A maximum t_1 evolution time of 20 ms was used. A mixing time of 50 ms was used to transfer the magnetization into the side chains. CP contact time was 1.75 ms.

5.3 Results

To identify the ET-1 signals in complex with its receptor, solution NMR experiments of unbound labelled ET-1 were conducted to assign the signals. Spectral assignments derived from solution NMR spectroscopic data was used as reference to correlate the solid-state NMR spectral assignments with these.

Two ET-1 preparations were analyzed by liquid NMR: (i) synthetic selectively labelled ^{13}CO (Met7, 12, Phe14), ^{15}N (Ile19, Ile20) ET-1 in buffer (pH 7.4), (ii) uniformly ^{13}C,^{15}N labelled ET-1 in acetic acid/water (pH 2.5). ET-1 was measured to obtain chemical shifts of the free ligand.

Two ET-1/ET_BR preparations were analyzed by solid-state NMR (ssNMR): (i) ^{13}CO (Met7, 12, Phe14), ^{15}N (Ile19, Ile20) ET-1 in complex with the detergent solubilized receptor (from *P. pastoris*), (ii) uniformly ^{13}C,^{15}N labelled ET-1 in complex with detergent solubilized receptor (from Sf9).

5.3.1 Solid state NMR study of synthetic ^{13}C, ^{15}N enriched ET-1 bound to the solubilized ET_B receptor from *P. pastoris*

Long-range internuclear distances (3-6 Å) provide useful complementary constraints for the bound ET-1 structure. Here, using ^{13}C,^{15}N labelled ET-1, intramolecular distance measurements should be performed in receptor sample by using REDOR (Gullion & Schaefer, 1989). The ET_B receptor was prepared with ^{13}C, ^{15}N- ET-1 as a complex (from *P. pastoris*) with a K_D for ET-1 of 10^{-12} M. Under such conditions ET-1 is bound to the receptor. Positions of the specifically labelled ^{13}C and ^{15}N were at carbonyl sites of Met7, Val12, Phe14 and at amide N of Ile19, 20, respectively.

Fig. 5.8 A shows the 1D ^{13}C solution NMR spectrum of labelled ET-1. The resonances from ^{13}CO (Met7, Val12, Phe14) ^{15}N (Ile19, Ile20) ET-1 in solution observed in a conventional solution experiment, however, give rise to four resonances at 175.7, 176.1, 177.0 and 177.5 ppm.

To study the solubilized ET_BR/ET-1 complex, a frozen preparation was used. For this purpose, the receptor complex was highly concentrated (~ 10 mg/ml) by ultracentrifugation

(see Chapter 3). NMR measurements were conducted in a 14.1 Tesla field at -13°C because of the improved signal-to-noise ratio at lower temperature. In addition, low temperatures should prevent loss of activity and change in the aggregation state (e. g. dissociation of the bound ligand) over an extended period of time. The challenge is the small volume of a typical 4 mm MAS rotor, i. e. 50 µl, and inherently the receptor complex is at low concentration (10 nmol, 0.5-0.6 mg). The 1D ^{13}C MAS spectrum of ^{13}C, ^{15}N ET-1 in complex with detergent solubilized ET_BR is shown in Fig. 5.8 B. The spectrum is dominated by the strong ^{13}C natural abundance signals arising from n-dodecyl-β-D-maltoside (10-40 ppm, 70-80 ppm and 100-110 ppm) present in the sample. The spectral intensity arising between 176 and 178 ppm has been assigned to the labelled carbonyl groups present in the peptide. Resonances (60-70 ppm) with low intensities were observed from natural abundance ^{13}C present from detergent-solubilized receptor. We attribute this to the low concentration of protein present in the sample with ^{13}C present at only natural abundance levels. The ^{13}CO resonances from the peptide have roughly about half the intensity of receptor backbone resonances, implying that the CO resonances are indeed from ^{13}CO labels rather than from the overlap of natural abundance carbonyl signals of the receptor. The individual CO resonances are rarely resolved. The 1D ^{15}N CP MAS experiment, while having lower sensitivity than the ^{13}C-detected experiment, showed no peaks (data not shown).

Unfortunately, from the spectroscopic point of view the CO labels have the disadvantage that the chemical shifts of carbonyl carbons have a very small chemical shift range, thereby making the distance measurement between carbonyl carbons extremely difficult. In terms of spectroscopic preference ^{13}C,^{15}N pairs between CO and N sites are the best choice. Based on the X-ray structure of ET-1 (PDB entry 1 EDN), the shortest intramolecular distances between ^{13}C,^{15}N labels are predicted to be 7.07 Å (^{13}CO-Phe, ^{15}N-Ile19) and 8.73 Å (^{13}CO-Phe, ^{15}N-Ile20), respectively. Such distances are too long to yield detectable dipolar couplings since specific ^{15}N to ^{13}C distances are measured on the basis of the r^{-3} distance dependence of dipolar coupling. The determination of distances (up to 5 Å) was hindered by three problems: low intensity, broad lines and the weak dipolar couplings resulting from both long distances between ^{13}C,^{15}N pairs and the low gyromagnetic ratio of ^{15}N.

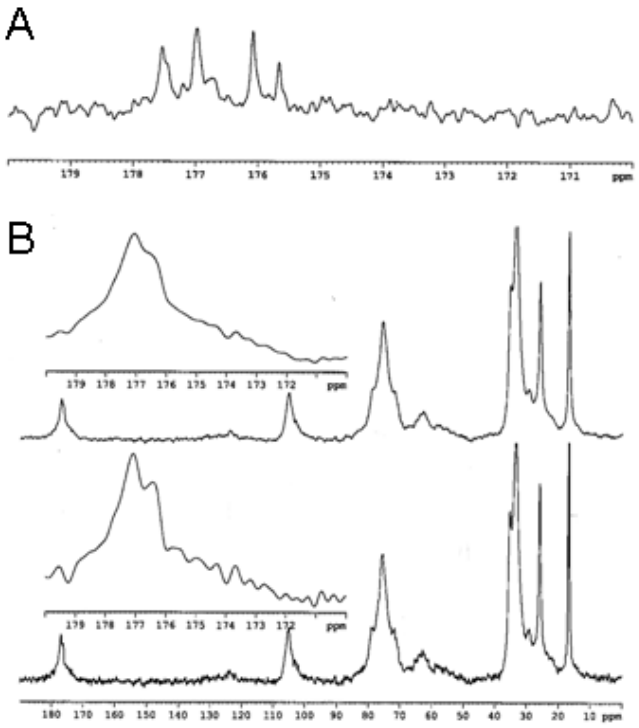

Fig 5.8. (A) 1D ^{13}C solution NMR spectrum of ^{13}CO (Met7, Val12, Phe14) ^{15}N (Ile19, Ile20) ET-1. (B) 1D ^{13}C CP/MAS spectra of the receptor complex. Insets show enlargements of the carbonyl region.

5.3.2 Solution NMR spectroscopy of recombinant uniformly labelled $^{13}C,^{15}N$-ET-1

To date no ^{13}C or ^{15}N assignments have been published for ET-1. Such an assignment in the solution state would allow a detailed comparison to be made with ^{13}C and ^{15}N assignments of ET-1 bound to the ET_B receptor obtained by solid-state NMR.

The NMR experiments were carried out in acidic aqueous solution both for solubility reasons and for comparison with previous NMR structural analyses of ET-1. NMR structural analysis of recombinant ET-1 in acidic solution revealed conformational preferences which are identical to those reported for synthetic ET-1 under identical experimental conditions (Dalgarno *et al*., 1992). The TOCSY spectrum (shown in Chapter 4) indicated that ET-1 adopts an ordered conformation under the conditions used. The peptide's amide nitrogen resonances are well dispersed in a ^{1}H-^{15}N HSQC spectrum (shown in Chapter 4). The assignments of the backbone (Cα, N and HN) were hence carried out based on the CBCA(CO)NH and HNCACB spectra. Owing to the well-resolved amide nitrogen resonances, the assignments from the two spectra were straightforward.

5.3.3 Solid state NMR study of recombinant uniformly ^{13}C, ^{15}N labelled ET-1 bound to the solubilized ET_B receptor from Sf9 insect cells

2D PDSD correlation spectroscopy measurements were carried out at 220 K in a 14.1 Tesla field on frozen solution prepared from equimolar mixtures of 4 mg receptor and 0.18 mg recombinant U-^{13}C,^{15}N labelled ET-1.

Amino acid type assignment can be made by direct correlation of the ^{13}C chemical shifts in the 2D ^{13}C homonuclear correlation spectrum. The mixing period for polarization transfer can be adjusted to probe connectivities between carbons separated by varying numbers of bonds. The homonuclear mixing time (50 ms) was optimized to provide Cα-Cβ and multiple bond correlations, based on monitoring the intensity of the Cβ region in 1D protein spectra as a function of spin diffusion mixing time. Under these conditions, the intensities observed arise almost solely from correlations between direct carbon carbon pairs, and interresidue cross peaks have still very low intensities, which allows to trace out the intraresidue connectivities for the different amino acid types in analogy to the J-correlation procedures in the solution state (Wüthrich, 1986).

We have compared the chemical shifts of the solid-state resonances with the assignment obtained on free ET-1 from solution NMR. The ^{13}C chemical shifts of all ET-1 residues from solution NMR are indicated in the 2D solid-state NMR spectrum of bound ET-1 (Fig. 5.9), allowing an almost complete assignment of the bound ET-1 Cα and Cβ cross peaks.

The ET-1 conformation is well defined in the receptor complex. Here the typical signal-to-noise ratio for a 10 d experiment can be seen. Fig. 5.9 is an expansion plot of the side chain region (0-80 ppm) of the carbon-carbon 2D PDSD correlation experiment recorded on U-^{13}C,^{15}N with a mixing time of 50 ms. Many signals are observed in the α-region (50-70 ppm). In addition, signals are observed in the β- and γ-region (0-50 ppm), due to magnetization exchange between Cα and labelled side-chain carbons. The spectrum reveals a variety of intraresidue side-chain correlations presented with horizontal and vertical bars to guide the eye through the recognition of the spin systems. The valine, leucines and isoleucines are easily recognized by the presence of the cross peaks between γ, β, and α resonances. Cross peaks which correspond to two-bond correlations are with the exception of these amino acids much weaker or not visible. Nearly all intra-residue cross peaks between Cα and Cβ can be found although the Phe, Tyr, Ile signals overlap. Missing or weak Cα and Cβ correlations are observed for Asp8, Asp18, Cys1, Cys3, Cys11 and Cys15, which were only ambiguously assigned by knowledge of secondary structure. Those of Cys1, Cys3 and Asp8 are shifted to lower fields. Residue Cys11 should have up-field shifted resonance which suggests that it is in

a stretch adopting a helical conformation. Towards the C-terminus a decrease in upfield shift of Asp18 should be observed. Cross peaks of the same amino acid type are not well resolved (Ile, Leu). There are two Ile and Leu residues in ET-1, thus resonance overlap is expected. The peaks of Ile19 overlaps with that of Ile20, which explains the particularly strong intensity of this resonance. Degenerate cross-peaks include Tyr13 (not resolved with Phe14, Ile19, Ile20) and presumably Cys1 (not resolved with Cys15). Secure assignments are only available for Ser, Ile, Val and Leu. Heteronuclear higher dimensional experiments will allow unambiguous assignments.

Many prominent cross-peaks corresponding to directly bonded pairs can also be seen in the CO-Cα region. Many weak two-bond peaks are particularly noticeable in the CO-Cβ region. In contrast to well-resolved aliphatic cross peaks, the CO-Cα region is fairly congested. The carbonyl region of this spectrum is not so informative as the Cα and Cβ region owing to the smaller carbonyl chemical shift dispersion. Significant overlap in the CO-Cα region proved to be very difficult in the verification of assignments, even when used in conjunction with the CO-Cβ region. One clear CO-Cβ cross peak is observed, e.g. a pair of bidirectional two-bond transfers (CO-Cβ and β-δ) is clearly detected for Ile side chains.

There are two stretches of amino acids that are apparently weak or broad due to structural heterogeneity: the first region includes residues Cys1-Lys9 and corresponds to the N-terminal loop region of ET-1, and the second region includes three residues of the helix (Cys15-Leu17). These residues from these regions are speculated to be weak because backbone motions cause poor performance in the polarization transfer sequences and the conformational heterogeneity of these regions could produce inhomogeneous broadening of the resonances. The solution NMR structure and structure-activity relationship studies strongly support the dynamics view. The conformation of the N-terminus was reported to be averaging on the NMR time scale, which is indicative of backbone motions. Whereas some of the published NMR models of ET-1 region Ser5-Asp8 adopt a β-turn conformation (Krystek *et al*., 1991; Tamaoki *et al*., 1991; Andersen *et al*., 1992; Dalgarno *et al*., 1992; Coles *et al*., 1994), others do not (Endo *et al*., 1989; Saudek *et al*., 1989, 1991; Aumelas *et al*., 1991; Munro *et al*., 1991; Reily *et al*., 1991). Asp8 and Asp18 are located before and at the end of the helix, which are known to be flexible in solution. This flexibility can result in a multitude of conformers, and such a mixture of conformers can lead to broad, nondetectable NMR signals in solid state. Lys9, also ill-defined in this stretch, is influenced by its less well-ordered neighbour Asp8. The C-terminal tripeptide Ile19-Trp21, is fairly well-defined in contrast to

Asp18. These observations are consistent with greater motion and correlate with the pH dependent effects for Asp8, Asp18, Glu10 and His16.

Despite the differences in temperature and pH nearly all of the solution chemical shift assignments coincide with cross peaks of the 2D MAS NMR spectrum except for the ambiguously assigned residues. Corresponding peak chemical shifts agree well at the two pHs, indicating that the pH difference does not induce a large peptide structural change.

Since chemical shifts are conformation-dependent, the marginal differences in chemical shifts suggest that the overall fold of ET-1 did not change upon binding. The amino acids for which the solid state NMR chemical shift deviates significantly from the corresponding solution values are Met7 and Trp21. The result predicts that the side chains of the residues Met7 and Trp21 of ET-1 are important to the ligand binding via a direct contact in the ligand-binding site or an indirect induced structural effect.

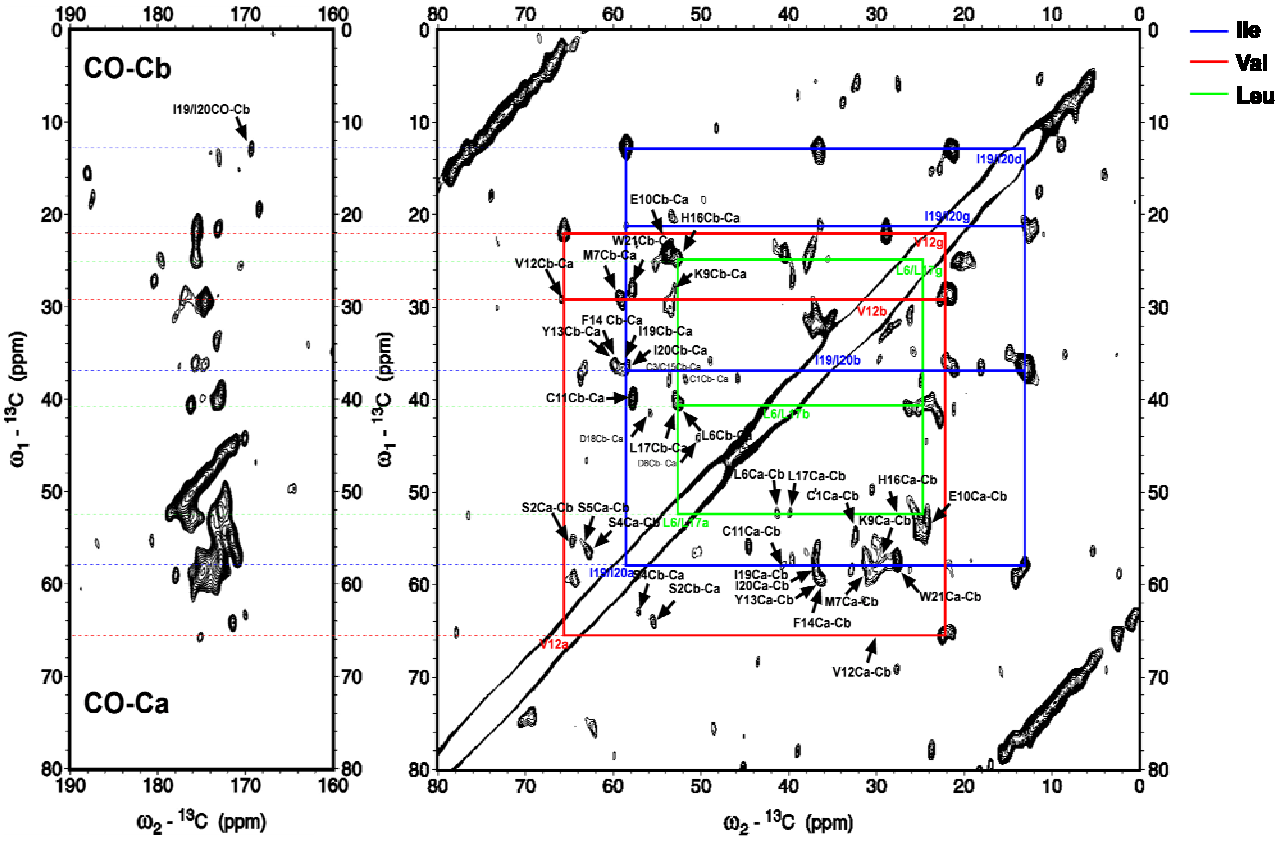

Fig. 5.9. 2D ^{13}C-^{13}C PDSD spectrum of ET_BR bound uniformly ^{13}C, ^{15}N-enriched ET-1 with a mixing time of 50 ms taken at 600 MHz. Lower case letters specify the type of cross peaks. Lines guide the eye for the connectivity paths.

5.4 Discussion

To be able to study the bound structure of ET-1 by solid-state MAS NMR, a previously developed methodology that is based on the detection of a large number of ^{13}C-^{13}C structure defining distances from several isotope-labelled samples, with minimal NMR experimental effort, was applied. The validity of this concept was previously shown on the globular SH3 domain of α-Spectrin in a microcrystalline state (Castellani *et al*., 2002) for which solution-state/solid-state NMR resonance assignments and a high-resolution X-ray structure are available. If a sufficient number of structural restraints are collected it is possible to calculate a structure. The applicability of this concept is primarily determined by protein availability and local structural or conformational homogeneity in the sample which influences the resolution of the spectra. This concept should be a new general approach to structural investigations of bacterial membrane proteins and receptor-bound peptides.

The preliminary data presented here demonstrates that it was possible through the incorporation of ^{13}C to selectively observe ET-1 whilst bound to the ET_BR by solid state NMR.

As a prelude to a full assignment of ET-1 bound to the receptor, an assignment of the backbone of free ET-1 under solution conditions was performed. These assignments as reference may be essential to further structural work, if there are restrictions regarding the concentration/ availability of functional ET_BR which affect ligand quantities that can be studied. Moreover, the chemical environment including detergent and receptor protein can hamper the unambiguous spectral identification of the bound ET-1 in a solid state NMR experiment.

5.4.1 Resonance assignment of U-^{13}C, ^{15}N-ET-1

Backbone resonances of ET-1 were assigned with manual assignment using sequential Cα, Cβ, N and HN chemical shift information derived from triple-resonance experiments (HNCACB, CBCA(CO)NH). To complete the assignment of the ^{13}C side-chain resonances, a 3D triple-resonance experiment referred to as C(CO)NH-TOCSY is still to be recorded with which the ^{13}C side-chain chemical shifts are detected in the third dimension of the 3D spectrum via a TOCSY-type transfer step. In addition, the assignment of the CO resonances will be achieved by using a 3D HNCO experiment, where the ^{1}H, ^{15}N frequency pair or each residue (i) is correlated with the CO shift of residue (i-1). Proton resonance assignments for ET-1 in solution have been published previously (Dalgarno *et al*., 1992).

5.4.2 Conclusions from chemical shifts and biological implication

2D solid-state NMR correlation data collected from detergent-solubilized ET_BR reconstituted with recombinantly prepared U-^{13}C, ^{15}N-ET-1 could be assigned the chemical shifts of the bound ET-1 with solution NMR chemical shifts. Previously several groups observed that there seems to be a general agreement between solution and solid-state NMR chemical shifts in proteins (Cole & Torchia, 1991; Detken *et al*., 2001; Luca *et al*., 2001). Secondary chemical Cα shifts of bound ET-1 observed for determined residues are remarkably similar with the shifts determined for free ET-1, although they were recorded at different temperatures, pHs and in different solvents.

The chemical shift depends on residue type and backbone conformation as well as on the nature and structural topology of the neighbouring residues. Conformational changes induce changes in chemical shift values, therefore, comparison to resonance assignments obtained on ET-1 in free form can be qualitatively interpreted in terms of conformational rearrangements due to ET-1-receptor interactions.

However, many of these chemical shift changes can also be attributed to effects of solvent and ionization state as a consequence of the experimental conditions. The receptor complex was measured in buffer at pH 7.4. ET-1 does not form a stable concentrated solution in water (Bennes *et al*., 1990). Therefore, solution NMR studies of free ET-1 were done in acetic acid/ water at pH 2.5 causing different ionization states of Asp8, Glu10, His16 and Asp18. These residues have possibly different hydrogen bonding potentials for interaction with the peptide backbone and side chains, which could have an important effect on the chemical shifts.

However, a large number of NMR conformational investigations of ET-1 and ET-1 analogues carried out with the aim of gaining insight into a possible bioactive conformation revealed that in aqueous solution and in organic systems ET-1 adopts similar conformational states. Comparisons between structures of ET-1 calculated from data acquired in different solvents are published (Coles *et al*., 1994; Lee *et al*., 1994; Wallace *et al*., 1995). A helix-like structure within ET-1 involving residues from K9-C15 and a turn involving residues S5-D8 were reported in both NMR and X-ray crystal studies. Great discrepancies were observed in the C-terminal hexapeptide tail as a result of its flexibility. While NMR results suggested that the hexapeptide tail folded back towards the bicyclic core, the crystal structure indicated that the K9-C15 helix was well defined and extended all the way through H16-W21. Since there is no solution structure data for ET-1 under conditions similar to those of the solid-state NMR experiments, it cannot be established if the pH is responsible for the observed change of

chemical shifts, but this remains a hypothesis. Here, it is speculated that binding may be responsible for the shifts reported here.

Assuming that the gross conformation of ET-1 at acetic and neutral pH is the same, the ET-1 signals of those carbons whose environments are changed by receptor binding will change position. The assigned backbone resonances of bound ET-1 at Met7 and Trp21 are shifted compared to that in free solution, indicating that the ligand possibly experiences a different electrostatic environment when in the receptor binding site. Noticeable is the change occurring in Met7 near the conserved Asp8-Lys9-Glu10 sequence. Lys9 is important for ligand binding, whereas position 10 plays a significant role in receptor activation (Miasiro *et al*., 1995). The loop region, where Met7 and where the greatest number of variable residues amongst the isoforms are located, has been implicated in the specificity of binding. Trp21 with a free C-terminal has been found to be a critical residue for ET-1 binding and activity. Its 3D orientation is important to receptor subtype selectivity, e. g. BQ-788 [2,6 dimethyl-piperidinocarbonyl-γ-MeLeu-D-Trp(Ninmethoxy-carbonyl)-D-Nle] possessed 1.2 nM affinity for the ET_B receptor with an ET_B/ET_A receptor selectivity ratio of greater than 200 (Ishikawa *et al*., 1994). If the helix 9-16 is continued up to residue 21 in the hypothetical receptor binding conformation of ET, as speculated by Tam *et al*. (1994) and found in the crystal structure (Janes *et al*., 1994), then all functional hydrophobic amino acids (Tyr13, Phe14, Leu17 and Trp21)) that are in either i+3 or i+4 relationship would lie on one side of the molecule.

Interestingly, the chemical shift data obtained in this solid-state NMR study suggests that bound ET-1 is predominantly non-helical in its C-terminal region. There is one line of evidence which correlates these shifts with nonhelical structure. The solid-state NMR results on ET-1 (16-21) Cα chemical shifts are consistent with non-helical structure observed for ET-1 by solution NMR, e. g. the Leu17 shift (52.9 ppm) is not within the one standard deviation (SD) range of 57.52 ± 1.23 ppm distribution of helical Leu Cα shifts in soluble proteins (Zhang *et al*., 2003). The Leu17 Cα shift fits well with both the 54.08 ± 1.31 ppm distribution of β-strand Leu Cα or with the 54.08 ± 1.31 ppm distribution of coil character Leu Cα, respectively. The solid-state Ile19/20 Cα shifts also correlate with nonhelical structure.

The side chain resonances of the N-terminal region (1-7) did not significantly change upon receptor binding with the exception of Met7, and on a chemical shift basis it appears that the N-terminal conformation is unchanged. The results are consistent with the proposal that the N-terminal residues of ET-1 (i. e. BQ3020) are not required for binding at the ET_BR, but the 8-21 region of ET-1. It cannot be confirmed that the entire C-terminus is helical in its

biologically active form due to absence of conformational rearrangement. Because most of the amino acids differences between ET-1, ET-2 and ET-3 occur within residues 2-7, this result supports the theory that the differences in ET_BR affinity between ETs can be attributed to the character and local position of its side-chains, rather than to a differently folded backbone. As side chain orientations are not fully defined in the various solution structures of the ETs, the possibility of differences in activity exists reflecting different side chain orientations, even though the global fold is similar.

However, much of this interpretation is somewhat speculative in that the interpretations depend upon this single spectrum. A full resonance assignment of at least the backbone atoms is a prerequisite for structural or dynamical investigations. The absence of a suitable assignment of the peptide backbone prevents any interpretation of the chemical shift data in terms of any preferred structure upon binding of ET-1 to the receptor. The spectrum presented here suggests that sufficient resolution is possible to allow a complete assignment of the ET-1.

5.4.3 REDOR distance measurement

The REDOR technique provides a means for measuring ET-1-induced structural changes within ET-1. The specific ^{13}C, ^{15}N labelling strategy was used to target distance measurements to the C-terminal tail to probe whether conformational change extends into it. Additional distance restraints should be deduced from the analysis of 1D MAS NMR REDOR between ^{13}C and ^{15}N labelled positions. Similar to that in the solution state, NMR structure determination in the solid state involves the calculation of families of molecular conformations that are consistent with the experimentally derived distance constraints. The number and precision of these determine the accuracy of the resulting 3D structure.

5.4.4 2D ^{13}C spin diffusion spectrum of [2-^{13}C] glycerol ET-1

To circumvent the need of a full assignment of all resonances, the preliminary assignments using solution NMR data can be verified by using the [2-^{13}C] glycerol ET-1 sample. In the spectrum of [2-^{13}C] glycerol ET-1 acquired with the same mixing time (50 ms), three types of residues, Ile, Val and Leu should be identified with relative ease. Isoleucines are labelled at the Cα, Cβ, Cγ1 and Cδ1 positions by the [2-^{13}C] glycerol labelling scheme (Fig. 5.7). Among these sites, Cγ1 has the lowest labelling level and Cβ the highest, thus cross peaks between Cδ1 and Cγ1 are not observed while cross peaks between Cδ1 and Cβ can be identified. The distinct upfield shift of Cδ1 should allow the assignment of peaks to the Cβ-Cδ1 couplings of isoleucine. Once the Cβ chemical shifts are identified, Cβ-Cα cross peaks are found by tracing

the connectivity paths. Valine Cα-Cβ peaks can be identified based on their strong intensities due to 100% labelling and the distinctly downfield shift of the Cα site. Leucine has 100% labelled Cβ-Cγ pairs, whose unique chemical shifts around (40 ppm, 25 ppm) allow it to be assigned straightforwardly. In addition to valine, several amino acids in the ET-1 sequence lend themselves to type-assignment by virtue of their distinct chemical shifts: leucine (Cβ and Cγ), isoleucine (Cδ1), and serine (Cβ).

5.4.5 Conclusions and outlook

Solution NMR chemical shifts were used to assign the bound ET-1 signals, especially as a preliminary to a more detailed study by solid-state NMR. Sequence-specific assignments would normally be determined by identifying sequential backbone connectivities, which are not performed yet. It is likely that some other sequence-specific assignments can be determined by comparison of solution and solid-state values. The good agreement of certain side chain shift patterns (e. g. Ile, Val, Leu) indicates that assignments based upon solution NMR chemical shift values are correct and a starting hypothesis for solid state NMR assignments. To independently proofread the solution NMR assignment and to obtain more structural constraints from the spectra, signals must be first assigned by using intra-and inter-residues spin connectivities acquired with a suite of solid-state NMR experiments. 2D ^{13}C-^{13}C experiments and the 2D planes from ^{15}N-^{13}C-^{13}C experiments permit the resolution and assignment of the side-chain resonances. These methods have already been used at high magnetic fields (750 MHz ^{1}H frequency) in studies of larger proteins, yielding a complete *de novo* assignment of α-spectrin SH3 domain (Pauli *et al.*, 2001). The results are encouraging, in particular the high signal-to-noise ratio of the 14.1 T 2D spectrum of a receptor complex sample containing only 80 nmol of protein. It is reasonable to expect that 2D ^{15}N-^{13}C correlation experiments are also viable. In combination with further ET-1/ET_BR preparations, which contain multiple labelled ET-1, the solid-state NMR derived distance restraints, obtained from spin diffusion buildup curves, should provide the basis for structure calculation of the receptor-bound ET-1. As demonstrated in the α-spectrin SH3 domain, the biochemical removal of nearest neighbour (^{1}H, ^{1}H) (Reif *et al.*, 2003) and (^{13}C, ^{13}C) (Castellani *et al.*, 2002) interactions simplifies the problem of detecting nontrivial structural constraints in polypeptides. On the other hand, spin dilution can complicate spectral assignment procedures, in particular in the presence of large detergent background signals and overall low signal to noise ratios. Simulated annealing procedures (Nilges *et al.*, 1988) and computational methods

that systematically search the entire conformational space will be used to define an ensemble of bound ET-1 structures consistent with the NMR measurements.

In general, the power of solid-state NMR for analyzing receptor-ligand complexes was previously shown by Luca *et al*. (2003).