Institut für Organische Chemie und Biochemie
der Technischen Universität München

Solution State NMR Investigations on the Structure and Function of Proteins: Application to Riboflavin Synthase

Vincent Truffault

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. St. J. Glaser

Prüfer der Dissertation:

1. Univ.-Prof. Dr. H. Kessler

2. Univ.-Prof. Dr. W. Hiller

Die Dissertation wurde am 20.12.2001 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 28.01.2002 angenommen.

Ce travail est dédié à mes parents, qui m’ont donné tous les moyens de développer mes aptitudes, et qui m’ont fait confiance lors de mon départ à l’étranger.

A Corinne, Eric, Anaï s et Maxime, qui m’ont soutenu et encouragé dans toutes les étapes difficiles.

Die vorliegende Arbeit wurde am Institut für Organische Chemie und Biochemie der Technischen Universität München in Garching unter Anleitung von Prof. Dr. Horst Kessler in der Zeit von Oktober 1998 bis Dezember 2001 durchgeführt.

Meinem akademischen Lehrer Prof. Dr. Horst Kessler möchte ich an dieser Stelle besonderes danken für das mir entgegengebrachte Vertrauen, sein Interesse an meiner Arbeit, seine uneingeschränkte Unterstützung in allen Belangen, die freie und ungezwungene Arbeitsatmosphäre und die hervorragenden Arbeitsbedingungen.

Mein weiterer Dank gilt:

-Prof. Dr. Dr. A. Bacher und seiner Gruppe für die fruchtbare Zusammenarbeit am RiSy-Projekt; dabei möchte ich vor allem Dr. Sabine Eberhardt und Dr. Holger Lüttgen für die unermüdliche Herstellung der zahlreichen RiSy-Proben und für die Kooperation bei den Bindungsstudien danken,

-Dr. Murray Coles, Dr. Tammo Diercks und Kerstin Abelmann für die großartige Zusammenarbeit am RiSy-Projekt, die die Basis für den Abschluß dieser Arbeit war,

-Dr. Gerd Gemmecker für die vielen lehrreichen Gespräche und Diskussionen sowie bei allen sonstigen spektroskopischen Problemen,

-meinen Praktikant Andreas Enthart, Nicolas Giraud, und Alexander Genest für Ihre selbständige und konstruktive Arbeit,

-nochmals Gustav Gemmecker, Michael John, Markus Heller und Melina Haupt für das intensive Korrekturlesen dieser Arbeit,

-Dr. Rainer Haeßner, Monika Goede und Beate Diaw für Rat und Tat bei den diversen Tücken des Computer- und Spektrometeralltags,

-Dr. Jens Liermann für die Hilfe bei Computerproblemen,

-den Sekretärinnen Frau Machule und Frau Bruckmaier für ihre professionelle Arbeit,

-allen weiteren Mitgliedern des Arbeitkreises für die zahlreiche Diskussionen und die angenehme Atmosphäre im Arbeitskreis,

-ma famille pour leur aide et soutient constant pendant mes études et la thèse.

Index

Abbreviations

amu	atomic mass unity
ATP	adenosine triphosphate
BMRB	biomagnetic research bank
COSY	correlation spectroscopy
CPD	composite pulse decoupling
CRINEPT	cross relaxation-enhanced polarization transfer
CRIPT	cross relaxation-induced polarization transfer
CSI	chemical shift index
CT	constant-time
HMQC	heteronuclear multi-quantum coherence
HSQC	heteronuclear single-quantum coherence
INEPT	Insensitive Nuclei Enhanced by Polarization Transfer
kDa	kilo-Dalton ($=10^3$ g/mol)
LED	longitudinal eddy current recovery delay
MQC	multiple quantum coherence
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
NMR	nuclear magnetic resonance
PASTA	protein assignment by threshold accepting
PDB	protein data bank
ppm	parts per million ($=10^{-6}$)
R_1	longitudinal/ spin-lattice relaxation rate
R_2	transversal/ spin-spin relaxation rate
RCSB	research collaboratory for structural bioinformatics
RiSy	riboflavin synthase
RMSD	root mean square deviation
SHIT	shared-incrementation time (semi-constant time)
T	absolute temperature
T_1	longitudinal/ spin-lattice relaxation time

T_2	transversal/ spin-spin relaxation time
TOCSY	total correlation spectroscopy
TPPI	time-proportional phase incrementation
TROSY	transverse relaxation optimized spectroscopy

1 Introduction and Aim of this Work

Proteins (Greek: proteios, primary) play central roles in all life processes, catalyzing biochemical reactions with remarkable specificity and serving as key structural elements in all cells and tissues. As enzymes, proteins catalyze both the digestion of food stuffs and the construction of new macromolecules. As collagen, actin, myosin and intermediate filaments, they control the structure and motion of cells and organisms. Antibodies protect us against disease; membrane proteins regulate ion transport and intercellular recognition. Repressors regulate gene expression, polymerases replicate genes, and histones help package DNA into chromosomes. Thus, proteins are involved in every aspect of life: catalysis, structure, motion, recognition, regulation [1].

Like the other biological macromolecules, the nucleic acids and the polysaccharides, proteins are polymers of smaller units. But unlike the nucleic acids, proteins do not have uniform, regular structures. This is because the 21 different amino acids from which proteins are made have widely differing chemical and physical properties [2].

In the last decade, nuclear magnetic resonance (NMR) spectroscopy has become a powerful technique, alternative to X-ray crystallography for the determination of macromolecular three-dimensional structures. Both techniques provide a detailed structure determination at atomic resolution.

The strength of X-ray crystallography lies in the possibility to study very large biological systems such as enzymatic complexes. In contrast, NMR spectroscopy is more limited by the size of the molecule to be analyzed. It is a younger technique and many efforts in this field over the last 10 years were directed towards the structure determination of proteins with larger molecular weight. It is actually possible to study biological complexes up to about 65kDa [3-6].

NMR has an advantage over crystallographic techniques in that experiments are performed in aqueous solution (close to physiological conditions) as opposed to a crystal lattice. In these particular conditions structure, flexibility and dynamics of proteins can be studied which is of prime importance for the probe of interactions

between proteins and other biological compounds. Structural studies to decipher protein-protein interactions, as well as protein-ligand and protein-nucleic acid interactions form the basis for the understanding of biochemical functions of proteins. Recent advances in the field of NMR spectroscopy, such as "SAR (*structure activity relationship*) by NMR" [7, 8] and screening [9] additionally underline the importance of NMR spectroscopy in this domain.

This work is concerned with the enzyme Riboflavin Synthase (RiSy), which catalyses the last step in the biosynthesis of riboflavin, performing the dismutation of 6,7-dimethyl-8-ribityllumazine. Gram-negative bacteria and certain yeasts are virtually unable to absorb riboflavin derivatives from their environment and are therefore dependent on the endogenous biosynthesis of the vitamin. The enzymes involved in the riboflavin pathway can therefore be considered as potential targets for the development of antibacterial and antifungal agents.

RiSy from *E. coli* consists of a 213 amino acids and is active as a 70 kDa homotrimer [10]. The RiSy monomer is composed of two domains, which are highly homologous and presumably have very similar structures. The 97 residue N-terminal domain, which contains the catalytic site, forms a homodimer in solution.

The aim of this work was to solve the solution structure of the N-terminal domain of RiSy from *E. coli* in the presence of bound riboflavin. As two ligand molecules are present per homodimer, the system is made up of four molecules overall. Experiments designed to observe exclusively inter-molecular contacts are specified. In particular, the strategy used for the elucidation of the binding site and the quaternary structure is outlined in detail (chapter 3).

The reaction of the riboflavin synthase is unusual in that it involves the transfer of a C_4 unit between two identical precursors, and despite considerable efforts over several decades, the mechanism of this dismutation reaction is not completely understood yet. Taken together with the presumed similarity in structure of the N- and C-terminal domains of RiSy, the dimerisation of the N-terminal domain and the knowledge of the binding site lead as a model for the structure of the RiSy

monomer, and allow a proposal of the arrangement of monomers in the RiSy trimer as well as of the reaction mechanism.

2 Protein Structure Determination by NMR-Spectroscopy

2.1 Preamble

The magnetic resonance phenomenon (discovered in 1946 [11, 12]) 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 = ½1, etc.) to give NMR spectra.

Each proton spin possesses a magnetic moment. When a sample is placed in a magnetic field B_0, its 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) [13, 14] 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.

Limitation by signal overlapping, can often be overcome by extending the measurements into a second dimension [15, 16]. All 2D NMR experiments comprise the same basic scheme. This consists 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.

The nature of the interaction during the mixing time depends on the type of experiment. Thus, in correlated spectroscopy (COSY [15]) and total correlated spectroscopy (TOCSY [17]) experiments the cross-peaks arise from through-bond scalar correlations, while in a nuclear Overhauser enhancement spectroscopy (NOESY [18, 19]) experiment they arise from through-space correlations.

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 ca. 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 from the H^N protons to the H_β 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 as shown in Tab. 2.1. 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 ca. > 70 amino acids .

Tab. 2.1: *Typical $^{3}J_{HNH\alpha}$-coupling constants in secondary structure elements of proteins* [20, 21].

Secondary Structure Element	$^{3}J_{HNH\alpha}$ [Hz]
Helices	< 6
α-helices	3.9
β-sheets	> 8
Anti-parallel β-sheets	8.9
Parallel β-sheets	9.7
Random coil	6-8

The assignment strategy proposed by Wüthrich [21] 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 [22, 23] to solve relative small protein structures of less than 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. However, the use of the NOE interaction for the sequential assignment can be dangerous, resulting in assignment errors because the NOE effect is a function of the secondary structure which is still not known at this stage of the study.

Problems associated with chemical shift overlap in 2D NMR spectra can sometimes be overcome by increasing the digital resolution and by using strong resolution enhancement digital filtering functions at the expense of sensitivity. Unfortunately, the gain resulting from these procedures is only minimal.

Another approach improves the spectral resolution by increasing the dimensionality of the spectrum to 3D [24, 25], while at the same time yielding important additional information about the system by combination of two 2D spectra, e.g., 3D soft COSY-COSY [26], soft NOESY-COSY [27], soft NOESY-HOHAHA [28].

The introduction of ^{15}N isotope labeling (which possesses a spin value of I = ½) of NMR samples permits to utilize scalar magnetization transfers such as $^{1}J/^{2}J$ scalar couplings that partially solve the sequential assignment problems and that exploit the heteronuclear frequency into a third dimension to avoid overlapping. The extension into a third and even a fourth dimension [29] permits to analyse interactions that could not have been extracted from the corresponding spectra. A 3D pulse scheme is easily designed by simply combining two 2D pulse sequences, leaving out the detection period of the first and the preparation pulse of the second. Thus, there are two evolution periods, t_1 and t_2, that are incremented independently, along with a detection period t_3.

The assignment strategy looks like the one proposed by Wüthrich, with the implementation of the third dimension to the traditional 2D spectra as a new point.

These strategies are schematically represented in Fig. 2.1. The parallel between the 2D-COSY/NOESY and their homologous 3D-^{15}N,^{1}H-HSQC-COSY/NOESY [30, 31] based information is also shown. Similar spectra based on the ^{13}C heteronucleus are exploited for the side-chain assignment. The success of the sequential assignment is however always dependent on the one of the H_α and H^N resonances and there is not really much of an increase in the molecular size studied by single ^{15}N labeled NMR samples.

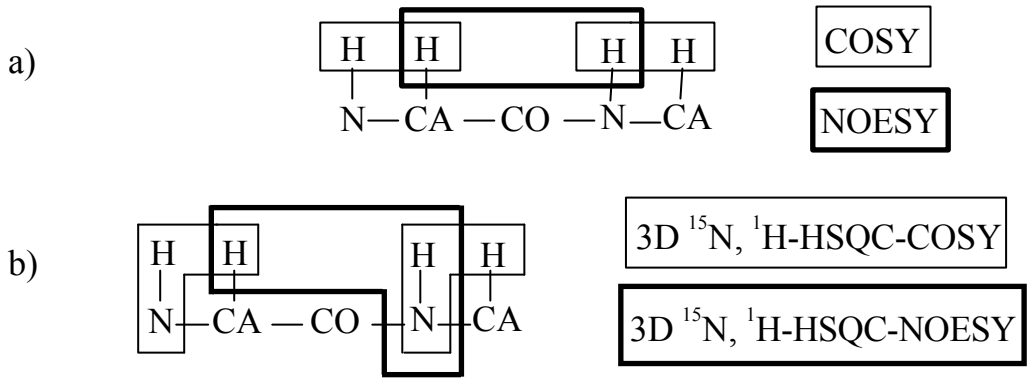


Fig. 2.1: *Assignment strategy: a) for unlabeled NMR sample; b) for single ^{15}N-labeled sample*.

A more useful approach suggests itself by employing uniformly, i.e.,up to about 95-98 % when cloned into overexpression systems [32, 33], ^{15}N and ^{13}C labelled proteins, thereby permitting access to the large heteronuclear couplings (Fig. 2.2) to transfer magnetization highly efficiently through bonds [29-31, 34].

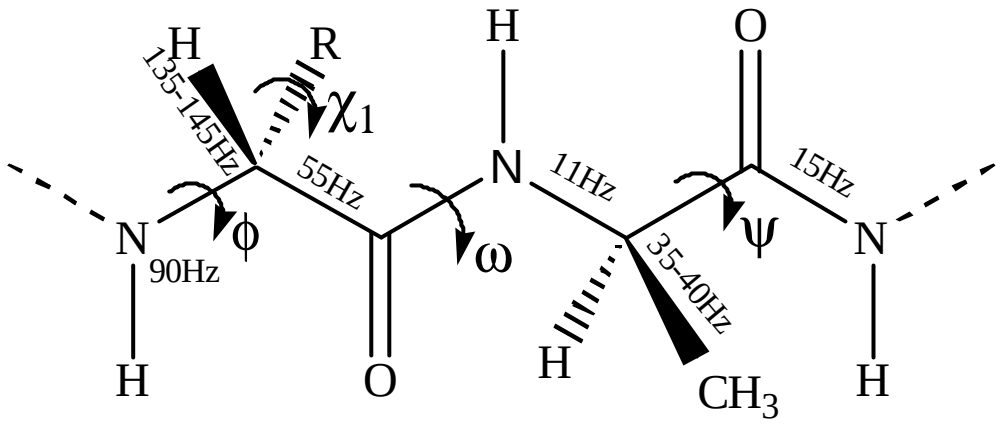


Fig. 2.2: *Two amino acids forming a polypeptide chain. Typical average* $^1J_{HX}$*-coupling constants for proteins as well as the principal angles are shown.*

However, isotopic labelling with ^{13}C results in a significant decrease in the proton transverse relaxation time T_2, making homonuclear J_{HH} correlation techniques even less sensitive. Therefore, separating homonuclear J-correlation spectra according to the chemical shift frequencies of attached ^{13}C nuclei, as previously used for smaller molecules, is inefficient for macromolecules.

Low sensitivity of hetero-atoms and above all the one of ^{15}N don't permit direct detection of the nitrogen signals. The experiments that will be mentioned in the following chapters are called inverse experiments and utilize the large proton sensitivity, i.e., they begin with ^1H excitation and finish with its detection. These experiments are also called "out-and-back" in comparison to their symmetrical architecture [35, 36]. The principle works for any system where high abundance, high gyromagnetic ratio nuclei are coupled to high abundance, low gyromagnetic ratio nuclei and employs a polarization transfer module.

Numerous successful procedures have been proposed for polarization transfer. For example, a well known of these, Overhauser polarization [37, 38] is applicable only when the relaxation of the low-sensitivity nuclei is dominated by interactions with a high sensitivity nuclear species. Also, the polarization transfer rate of this method is usually quite low for large macromolecules.

In contrast, in the INEPT [35] pulse sequence the modulating effects of heteronuclear coupling and chemical shift are separated by the use of refocusing pulses to achieve the coherence transfer. Fig. 2.3 and Fig. 2.4 display some typical pulse sequences where the desired combination of heterouclear coupling and chemical shift can be realized. The INEPT transfer is generally used as standard magnetization transfer between two heteronuclei, but for very large molecules, it is limited by efficient cross-correlated relaxation rates.

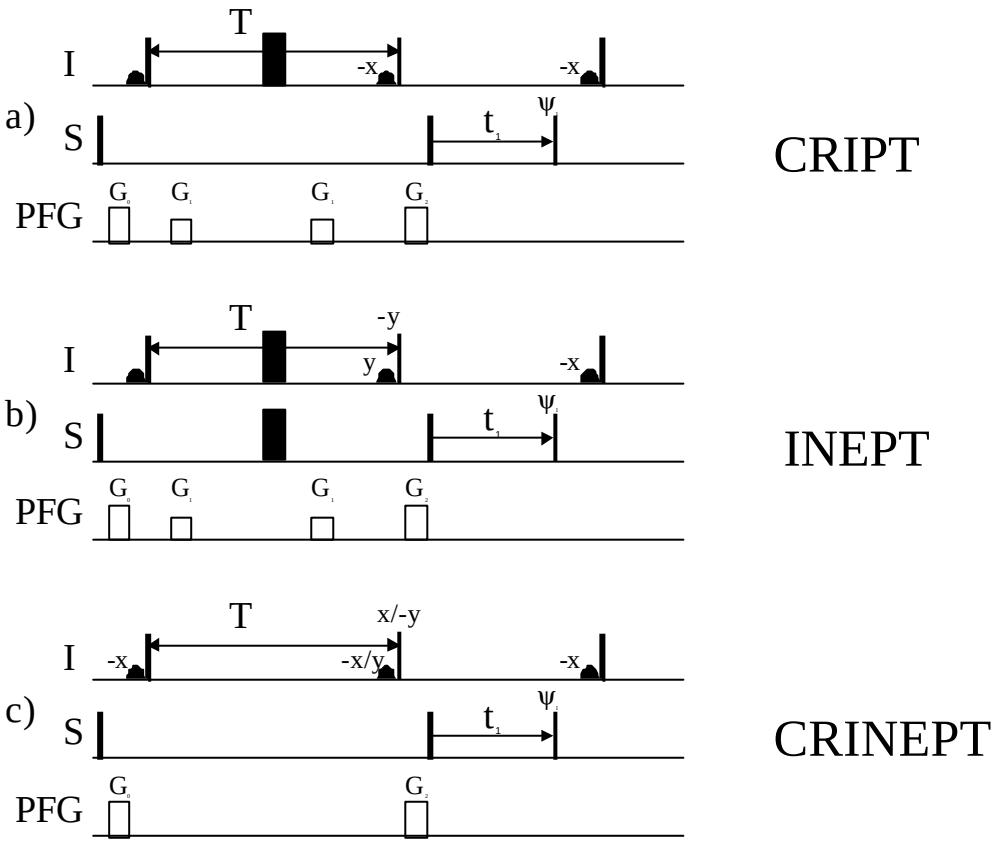


Fig. 2.3: *Polarization transfer pulse sequences: a) Cross relaxation-induced polarization transfer (CRIPT), b) INEPT, c) Cross relaxation-enhanced polarization transfer (CRINEPT). In all sequences, water flip-back on +z is achieved by 90° selective pulses. The pulse phases are x if not precised.* $\boldsymbol{Y}_1 = (x, -x)$.

As the rate of the cross-correlated relaxation is proportional to the correlation time τ_c (which means proportional to the size of the protein), the rate of the polarization transfer via this mechanism becomes efficient enough for solution NMR with very high molecular weight (≥ 100 kDa).

During the CRIPT [39] sequence, there is a transfer from in-phase ^{1}H-magnetization to anti-phase ^{15}N-magnetization via cross-correlated relaxation. The ^{1}H chemical shifts are refocused by the 180° pulse as well as the direct $^{1}J_{HN}$ coupling. During the CRINEPT [40], the absence of 180° pulse results in magnetization transfer by both cross-correlated relaxation (CRIPT effect) and scalar coupling (INEPT effect). Also, the CRINEPT sequence takes the advantages of both, the INEPT and the CRIPT effects. For short τ_c, the cross-correlated relaxation rate (R_c) is small and only INEPT polarization type contributes to CRINEPT; for longer τ_c, R_c becomes large and CRIPT transfer type dominates the polarization transfer of the CRINEPT pulse sequence.

Otherwise, to reduce the intensity losses due to relaxation during the evolution time, there are some possibilities for concatenation of an evolution period and a detection period, known as constant-time (CT) [41-43] and semi-*constant time* or *shared-incrementation time* (SHIT) [44, 45]. Both are represented in Fig. 2.4.

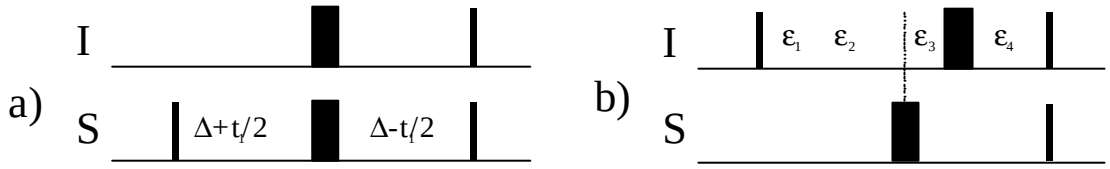


Fig. 2.4: *Two relaxation optimized periods: a) Constant-time evolution period (CT); b) Shared-incrementation evolution period (SHIT).*

High resolution in the ^{13}C spectrum can be obtained in an efficient manner by using a *constant time* evolution period. The *constant time* is usually chosen for coherence transfer delays exceeding the desired acquisition time, i.e., $1/(2\ ^{1}J_{XY}) > t_1$ max where $^{1}J_{XY}$ represents the one bond coupling constant between the nuclei X and Y. The uniformity of one-bond $^{1}J_{CC}$ couplings makes it possible to optimize the CT ^{13}C period offering high resolution at a minimal cost in sensitivity. Thus, during CT, ^{13}C homonuclear decoupling and ^{13}C single quantum coherence evolve.

The *Shared* (semi *constant-time*) evolution time is chosen for $1/(2\ ^{1}J_{XY}) < t_1$ max. In contrast to the CT period, SHIT evolution time is used for large coupling

constants that have to be refocused during the evolution period. Then, $\Omega_I + J_{IS}$ evolve during SHIT in comparison with $\Omega_S + J_{IS}$ during CT. Due to the small transfer delay, SHIT combines the refocusing of the couplings with an incrementation period to have enough resolution.

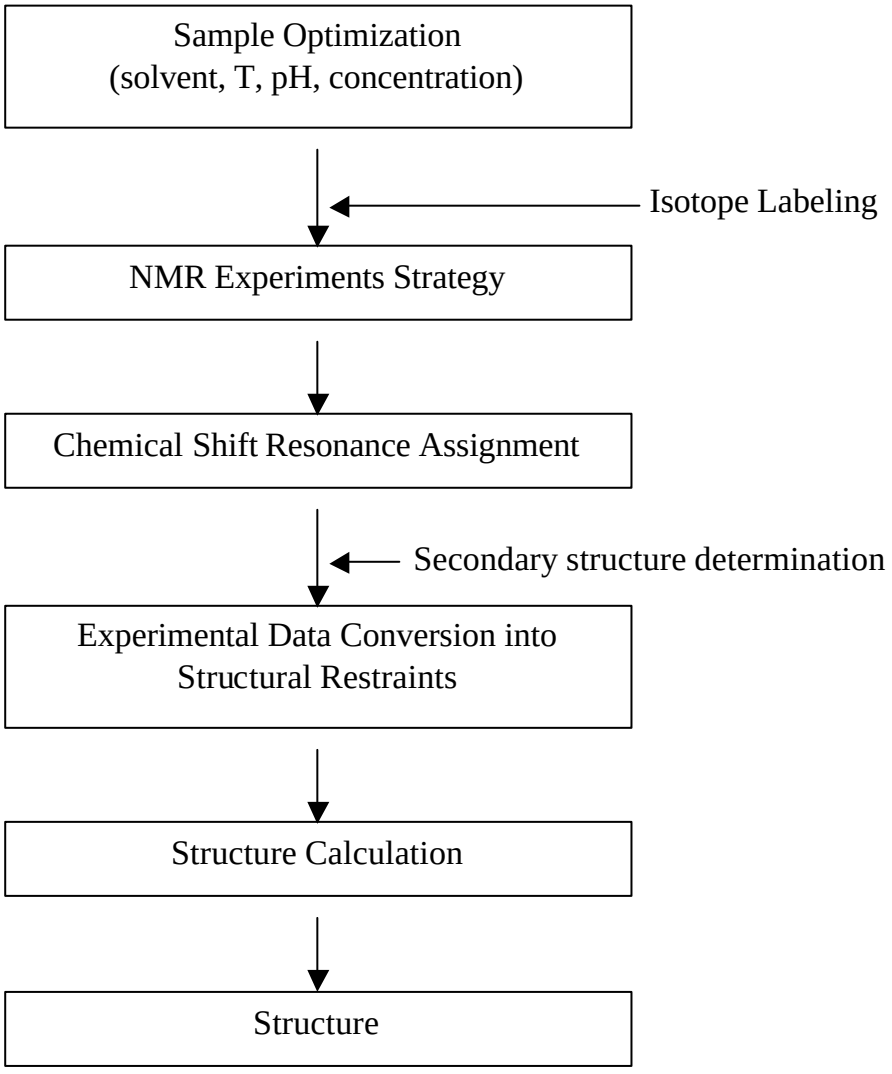
To go further into the study of larger biological complexes by NMR and to conclude, partial deuteration [46] of the samples yields an attenuation of the relaxation rates ($\gamma_D \sim 1/6.5\ \gamma_H$); selective isotope labeling of some backbone residues [46-48] results in fewer signals in the spectra and thus a reduction of signal overlapping. TROSY/CRINEPT [3, 49, 50], 4D experiments [29] as well as technical developments in the field of NMR spectroscopy (higher magnetic fields, cryoprobes, gradients, etc.) all allow the structure determination of larger proteins. Also new approaches yield structural restraints that are orientational in complement to the standard distance restraints. Anisotropic spin interactions such as residual dipolar couplings [51] (RDC) and chemical shift anisotropy [52, 53] (CSA) provide useful structural information that is not based on NOEs. To date, RDC is generally used together with NOEs, supplying the lack of long range NOEs by other structural restraints. However, in the last months a study was reported where the structure of a protein backbone was determined using only RDCs [54], suggesting the possibility to solve protein structures without any distance information and thus opening new perspectives for the complete structure determination of proteins larger than 100 kDa. The actual size limit for the study of proteins lies at ca. 65 kD [3, 5, 6].

2.2 Structure Determination of [U-^{15}N, ^{13}C] Labeled Proteins

2.2.1 Introduction

As mentioned in the latter chapter, the problems of limited resolution among the H_α resonances and the conformation dependence of the NOE and ^{3}J coupling constants are overcome in the assignment strategies for ^{13}C-, ^{15}N-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 labeled proteins is shown in Tab. 2.2.

Tab. 2.2: *General strategy for the protein structure determination with NMR spectroscopy***.**



To succeed in this whole procedure 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 the assignment of ^{13}C-, and ^{15}N-labeled proteins are summarized and the corresponding assignment strategies are discussed.

2.2.2 Assignment of the backbone chemical shifts

One of the first strategy based on triple resonance experiments was proposed by Bax's group [55] and employs several 3D experiments (here not shown) to correlate the resonances of the peptide backbone (H^N(i), N^H(i), C_α(i), H_α(i), C_α(i-1), H_α(i-1), C'(i) and C'(i-1)).

Almost experiments contain the ^{15}N- and H^N-resonances, therefore allowing the use of this pair of spins as reference and starting point for further assignment of other resonances. The HNCA [56] experiment, for example, correlates the H^N and ^{15}N chemical shifts of residue (i) with the $^{13}C_\alpha$ shifts of residue (i) (via $^1J_{NC\alpha} \approx 7\text{-}12$ Hz) and residue (i-1) (via $^2J_{NC\alpha} < 8$ Hz), thereby providing sequential connectivity information. A complementary experiment to the HNCA, the HN(CO)CA [57] correlates, in contrast, the amide proton and nitrogen resonances of an amino acid only with the C_α 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 $^{13}C_\alpha$ via the intermediate carbonyl nucleus. Generally, the intensity of the C_α (i) and C_α (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 H^N-N-C_α correlations in the HNCA. An HNHA [58] experiment completes the assignment of

the α-resonances. The spins $H_\alpha(i)$ are correlated to the H^N via a $^3J_{HNH\alpha}(i)$ coupling that connects only the resonances of the same residue.

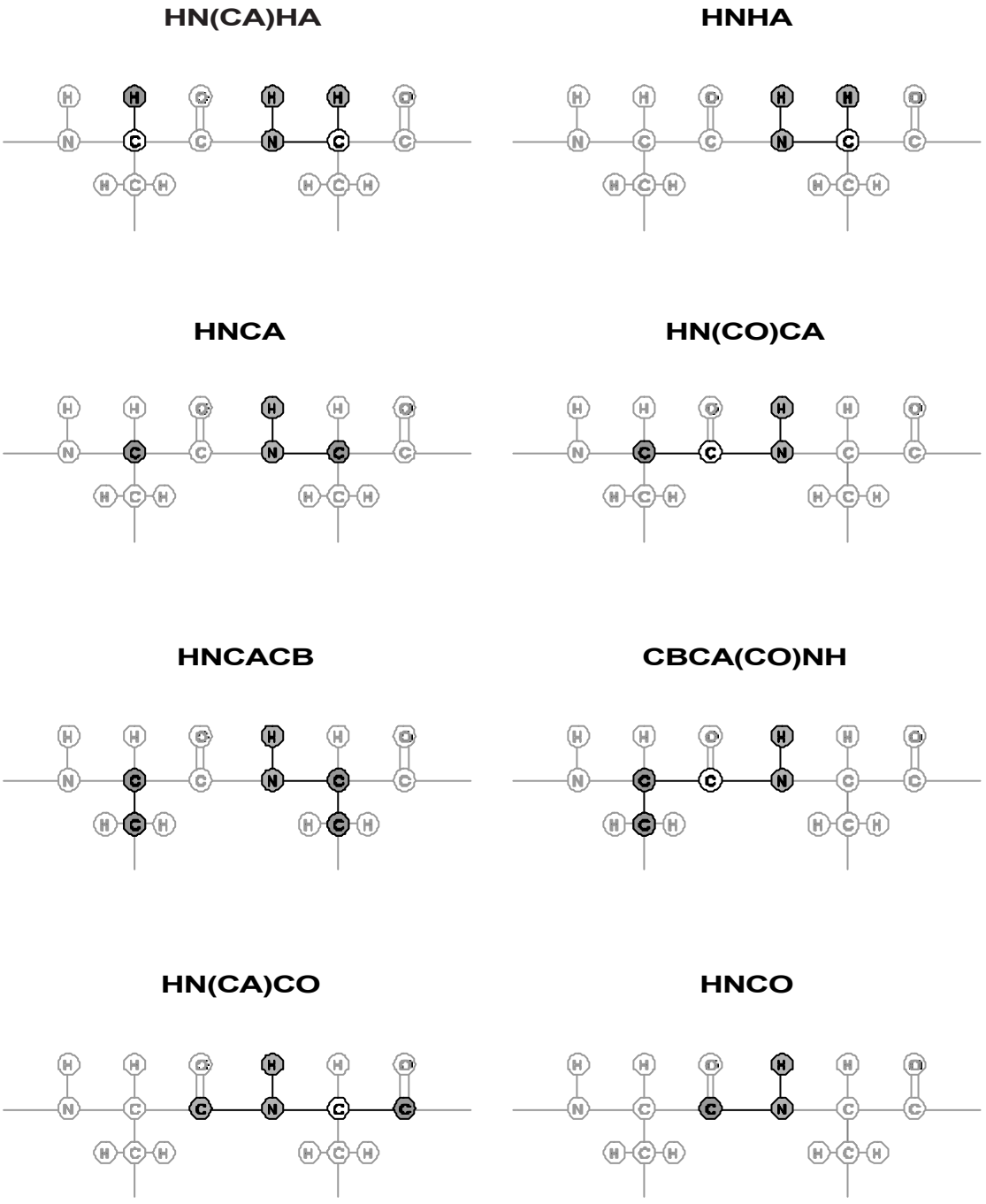


Fig. 2.5*: Combination of triple resonance experiments for the backbone assignment of doubly ^{15}N-^{13}C labeled proteins.*

To circumvent the severe overlap of the α-resonances, further developments involved the chemical shifts of side-chain carbon and proton spins (especially C_β and H_β) to achieve the sequential assignment. It follows then the recommended set of experiments (Fig. 2.5): HNCO [56], HN(CA)CO, HNCA [56], HN(CO)CA [57], HNCACB [59], CBCA(CO)NH [60], HNHA [58], HN(CA)HA [61, 62], HNHB [63].

Among these atoms, C_α and C_β are of prime importance, first because their chemical shifts show a large spectral dispersion ($C_\alpha \approx 25$ ppm; $C_\beta \approx 60$ ppm), and then because taken together, these shifts are characteristic for the identification of the amino acids (Fig. 2.6).

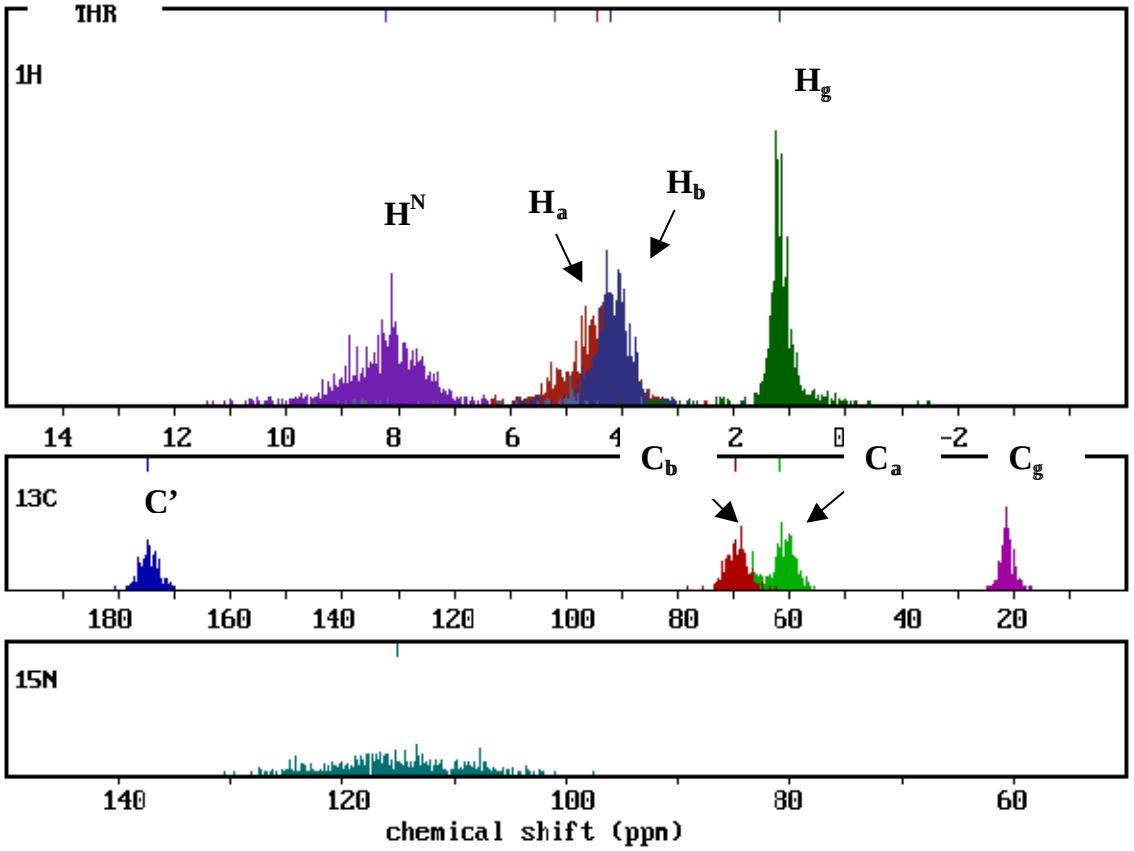


Fig. 2.6: *Protein threonine chemical shift distribution histogram extracted from the BMRB data base (http://www.bmrb.wisc.edu/search/amino_acid_cs_dist_plots.html). The typical threonine **a**/**b**-carbon pattern is shown as well as all the other resonances.*

The most significant experiment for the assignment of the backbone resonances is the HNCACB. This experiment yields the C_β shifts (in position (i) and (i-1)) in addition to those coming from the HNCA. The C_α and C_β correlations have opposite signs and can thus be distinguished (Fig. 2.7). 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.

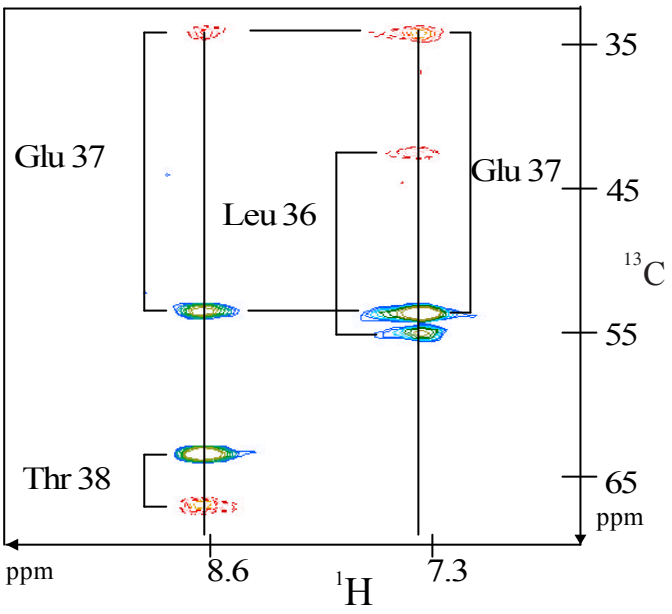


Fig. 2.7: *Two spin system traces showing the four aliphatic carbon correlations in a HNCACB. The C_α(i) and C_α(i-1) are colored in blue; the C_β(i) and C_β(i-1) are colored in red.*

Protein structure determination with NMR is very time consuming. In order to facilitate it and to make it faster and in line with this strategy, a large set of computer programs [64-67] (and references therein) was developed. To date, these programs offer a semi-automatic approach very helpful for the sequential assignment.

The HNCACB allows a complete assignment of the C_α and C_β resonances experiment, providing together the recognition of the residue and sequential information, which exactly satisfies a fast and unambiguous manual and/or computational automated approach. To complete the β-resonances assignment, an

HNHB [63] gives exclusively the H_β shifts via a $^3J_{NH\beta}$ coupling, and similarly to the pair HNCA/HN(CO)CA, the CBCA(CO)NH provides specifically the sequential information, i.e., the C_α/C_β (i-1) only.

Proline resonances that don't possess any amid proton are only indirectly assign via the experiments mentioned above; by the (i-1) information derived from the CBCA(CO)NH, for example. If several prolines are successively positioned in the primary sequence, the attribution of their chemical shifts will only succeed by the use of specific experiments [68, 69].

2.2.3 Assignment of the side-chains chemical shifts

Assignment of the side-chain resonances and especially of the proton chemical shifts is a precondition for analyzing NOE interactions which yield important distance restraints. For example, valine and leucine side-chains are often situated in the hydrophobic core of the protein thus giving a lot of distance constraints resulting in structural information. The success of this assignment is then directly related to the structure quality.

HCCH-TOCSY [70], C(CCO)NH-TOCSY [71] and H(CCCO)NH-TOCSY [72] represent the classical set of experiments involved in the assignment of the side-chain resonances. Both, carbon and proton shifts are observed in the HCCH-TOCSY but their assignment stays somewhat difficult due to severe signal overlap. Also this experiment will often be evaluated in combination with the C(CCO)NH-TOCSY and the H(CCCO)NH-TOCSY. The former one correlates the carbon shifts of the whole side-chain spin system with the $^{15}N/H^N$ (i+1) backbone shifts, while the latter one correlates their proton shifts with the successive amide group. However the efficiency of the TOCSY sequence decreases with the length of the side-chain; the magnetization is dispersed along the side-chain and only a part will be transferred to the C_α. Moreover, these experiments become insensitive with increasing molecular weight due to fast relaxation of the aliphatic carbons. Also for proteins larger than 15 kDa, an alternative is to use the HCCH-COSY [73] in supplement to the HCCH-TOCSY. These experiments can be acquired either in a

3D or a 4D manner. For the 3D version, it will be preferred to acquire the spin system with a H(C)CH-TOCSY (thus observing the whole spin system via the protons) while the trace of the side-chains is followed by the (H)CCH-COSY. The trace of the spin system is better followed by the carbon shifts due to a larger chemical shift dispersion of the heteronuclei, and the $^{1}J_{CC}$ is more sensitive than the $^{3}J_{HH}$ that results in a more efficient COSY transfer. The $H_{\alpha/\beta}$ and $C_{\alpha/\beta}$ chemical shifts of the side-chains (generally already known from the backbone assignment) are then used to link the side-chain spin systems to the backbone shifts.

As mentioned above the assignment of the side-chain resonances, which means also their stereospecific assignment, significantly improves the quality of protein structures determined by NMR. Both the number and the accuracy of distance constraints involving diastereotopic β-/γ-methylene groups of Val and Leu or the primary amide groups of Asn and Gln, for example, increase the definition of the tertiary structure. This kind of information can be obtained from a combination of vicinal coupling constants (e.g., the $^{3}J_{NH\beta}$ coupling constant measured from the HNHB [74] experiment gives access to the values of the χ_1 angle) and intraresidual NOEs [75]. Each of the three possible rotamers can be then identified as shown in Fig. 2.8.

Several methods were developed in order to allow the stereospecific identification of diastereospecific atoms or groups. For example, the diastereotopic methyl groups of Val and Leu can be discriminated using biosynthetically directed fractional ^{13}C-labeling (10-20 %) [76, 77]. Biosynthesis of the amino acids valine and leucine is known to be stereoselective [78, 79]. The stereospecific distinction between the pairs of methyl groups is then clearly observed in standard high-resolution ^{13}C-HSQC spectra, where the ^{13}C component of the pro-R methyl group is a doublet (splitted by the $^{1}J_{CC}$ coupling) while the ^{13}C NMR signal of the pro-S methyl group is a singlet. Other strategies make use in the same way of the incorporation of stereoselectively deuterated amino acids [80, 81].

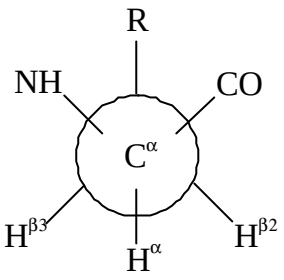
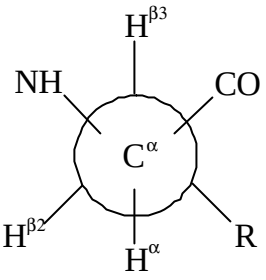
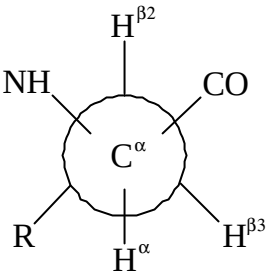
χ_1	60°	180°	-60°
			
$^3J_{\alpha\beta2}$ [Hz]	< 4	< 4	> 10
$^3J_{\alpha\beta3}$ [Hz]	< 4	> 10	< 4
NOEs	α-β2 ≈ α-β3	α-β2 > α-β3	α-β2 < α-β3
NOEs	NH-β2 < NH-β3	NH-β2 ≈ NH-β3	NH-β2 > NH-β3

Fig. 2.8: *Scheme for obtaining stereospecific assignments of b-methylene protons on the basis of characteristic coupling constants and NOE patterns.*

A similar approach is, of course, not applicable for the individual assignment of exchanging nuclei such as side-chain amide protons of asparagines and glutamine residues. However, several methods have been introduced to assign these specific polar and charged side-chains spin systems as well as those of Asp and Glu [82, 83]. The H_2NCO-E.COSY [84], for example is based on a triple resonance backbone experiment, the HNCO, and gives E.COSY- [85] like multiplets from which $^3J_{H\delta C\beta}$ and $^3J_{H\varepsilon C\gamma}$ coupling constants can be determined for asparagine and glutamine side-chains, respectively. The two amide protons occur in the Z (*syn* to the carbonyl oxygen) or the E position, respectively, which results in a large and small coupling constant and thus in an unambiguous assignment. Similarly, Bystrov [20] remarks that the $^1J_{NH_Z}$ (~ 90.2 ± 0.9 Hz) is smaller than the $^1J_{NH_E}$ (~ 93.2 ± 1.3 Hz), thus easily discriminating both ^1H by measurement of their coupling constants [86].

When the stereospecific assignment is not possible, a set of pseudoatoms replacing the diastereotopic hydrogen atoms can be introduced [87]. This is however

only a compromise, since the use of these pseudoatoms reduces the precision of the experimental conformational constraints.

Finally, as mentioned in the preceding chapter, the strategy permits also a computational automatic or semi-automatic approach for the chemical shift assignment in general, as remarked by Kay's group: "Because of the low level of resonance overlap in the 3D spectra, much of the 3D peak picking can be done in a fully automated manner. ….. thus dramatically reducing the amount of human intelligence (!) and labor required in the assignment process." [56]

2.2.4 Secondary structure determination

The secondary structure represents local spatial arrangements of frequently found conformations and can be specified once the chemical shift assignment is done. The most commonly used method to determine the position and the type of the secondary structure elements is based on the study of the chemical shift values. A particular attention was always attributed to these values, based on the assumption that it should be possible to achieve a complete solution structure determination only via the chemical shift parameter. In addition to its role played in the determination of the secondary structure [88, 89], the chemical shift is involved in the NMR studies as for example in monitoring folding transitions [90], in quantifying main chain flexibility [91], in refining tertiary structure [92], and in the definition of the binding site through chemical shift differences between apo- and holo-forms (see chapter 3). Thus, the chemical shift information is often applied to solve conformational problems of biological significance.

In addition to H_α, the most affected shifts are C_α, C_β and C' [93]. Statistical lists were established to define random coil chemical shifts [88, 94], also called primary shifts. The chemical shift difference between its experimental value and its random coil value is called secondary shift. This secondary information is used for identifying regions of secondary structure. It is possible to implement this method into an algorithm, where the secondary motives are classified in three different categories (Tab. 2.3), defining thus the chemical shift index [88, 89] (CSI).

Tab. 2.3: *Sign of the secondary shifts for C_α, C_β, C' and H_α in different secondary structures. Those whose secondary shifts stay in the range of the random coil values don't correspond to a secondary structure element.*

	C_α	C_β	C'	H_α
α-helices	> 0	< 0	> 0	< 0
β-sheets	< 0	> 0	< 0	> 0

The presence of a particular secondary structure element is identified by the consecutive repetition of at least three amino acids for which the sign of the CSI stays constant. However, even if the motives of secondary structure are well predicted by the CSI, their exact position along the primary sequence is difficult to establish.

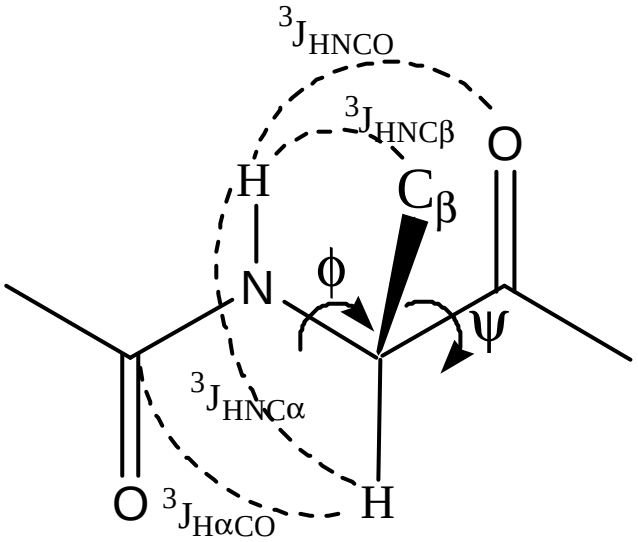


Fig. 2.9: *Schematic representation of the vicinal $^1H,^1H$- and $^1H,^{13}C$-couplings measured in different experiments for the determination of the f-/y-angles.*

A more precise delineation can be achieved from the measurement of the $^3J_{HN}{}_{H\alpha}$ coupling constants (directly correlated to the values of the ϕ angles [20, 95, 96]; see Tab. 2.1), since a secondary element is defined by a specific combination of

the backbone ϕ- and ψ-angles [97]. While several experiments allow the measurement of the ϕ-angle values [58, 98], access to the J-couplings that define the ψ-angles is limited by the presence of the oxygen and nitrogen heteronuclei bound to the C' (Fig. 2.9). As a result, only the $^{3}J_{HN_{CO}}$ provides structural information for the ψ angle.

Finally, a complete determination of the topology will occur after close examination of strong and medium backbone NOE contacts, which show typical NOE patterns [99] and allows thus the discrimination between parallel and anti-parallel arrangement of the β-sheets. These NOE patterns as well as characteristic combinations of the ϕ- and ψ-angles are shown in Tab. 2.4.

Tab. 2.4: *Typical angles and NOE patterns for the determination of secondary structure elements in proteins.*

angle/distance	α-helix	β-sheet
ϕ	-57°	-139°
ψ	-60°	135°
$d_{NN}(i,i+1)$	strong	weak
$d_{\alpha N}(i,i+1)$	weak	very strong
$d_{\alpha N}(i,i+3)$	medium	-
$d_{\beta N}(i,i+1)$	medium	-
d_{NN}(cross strand)	-	weak
$d_{\alpha N}$(cross-strand)	-	very strong
$d_{\alpha\alpha}$(cross-strand)	-	very strong in anti-parallel β-sheets

2.2.5 Tertiary structure determination

The finding of a tertiary structure results in the success of all the steps discussed in the preceding chapters. The last stage in the long procedure of protein structure determination with NMR spectroscopy consists in the collection of tertiary structural information. To date, several techniques with different levels of development offer the possibility to access this information.

The classical way to determine the tertiary structure of biomolecules is based on interproton distances r which are derived from NOEs. Perturbing the populations of stationary states within a dipolar coupled spin system induces changes in the intensities of signals via the NOE. Intensity changes result from cross-relaxation when (in a system of two nuclei) the double-quantum relaxation (W_2) differs from the zero-quantum relaxation (W_0).

The cross-relaxation rate is proportional to the inverse sixth power of the interproton distance (r^{-6}). Thus, if one interproton distance is known, all the other interproton distances can be easily calculated. However, only the initial build-up rate of the NOE is proportional to r^{-6} and can result in exact quantification of cross-peak intensities. This initial rate approximation is satisfied by the use of short NOE mixing times, giving rise to low signal-to-noise ratios of the cross-peak intensities. Also, in most applications, higher signal intensities are obtained by using longer NOE mixing times. As a consequence, spin-diffusion contributes to the NOESY cross-peak intensity which is no more directly proportional to the cross-relaxation rates. Thus, instead of giving precise values, interproton distances are often empirically classified into three or four categories, strong, medium, weak and sometimes very weak corresponding in constraints on upper distances of 2.7, 3.2, 4.0 and 5.0 Å, respectively.

In the strategy proposed by Wüthrich [21], it is shown that 2D-homonuclear NOESY experiments are enough to solve the structure of proteins up to 70 amino-acids. For larger biological complexes, increasing signal overlap can be overcome by the use of singly and double labeled samples which allows the possibility to extend this technique to 3D or 4D heteronuclear experiments. One of the first heteronuclear edited NOESY experiments proposed in a 3D manner was the HNH-NOESY [26, 27], that provides correlations between the NH amide spin pair and all the other protons closer than about 5 Å. The side-chain/side-chain contacts (especially the one forming the hydrophobic core of the proteins) can be specifically observed via an HCH-NOESY experiment [31, 100] for example. Problems of intensive signal overlapping in the region of the aliphatic protons can be solved using a combination of NNH- [101], CCH-, NCH-, and CNH-NOESY [102]

experiments that exploit the large spectral dispersion of the heteronuclei (see chap. 2.1; 2.2.2). Grzesiek *et al*. proposed also a 4D HNNH-NOESY [103]. However, to our experience the diagonal free 3D CNH-NOESY [102] yields the best dispersion and highest resolution at the compromise of experiment time and cross-peak intensity.

The fold of the protein can thus be determined by identification of NOEs between residues far apart in the sequence but close together in space. This step can only be accomplished once complete (or almost complete) resonance assignment is available. An average of 13-15 NOEs per residue is estimated representing a high quality structure.

New techniques emerged in the last years and their application is possible once the nitrogens and their attached amide protons are assigned. The origin of the phenomenon, well known in the liquid crystal community [104] and in solid-state NMR, is based on the anisotropic magnetic interactions and is subsequently normally not observed in solution NMR. However, dissolved in media such as phospholipid bicelles [105, 106], filamentous phages [107, 108] and others [109-111], the proteins are partially oriented with respect to the principal magnetic field and the residual dipolar couplings [112] and/or chemical shift anisotropy [52, 113] can be measured.

This class of experiments yields structural restraints that are orientational rather than distance based and are available very early in the structure determination process: at completeness of the backbone assignment. Also in combination with other methods like perdeuteration of non-exchangeable protons [114] and/or the use of the TROSY scheme, the assignment of the amide spin pair can be extended to very large biological complexes and the “long and laborious” process of solution structure determination with NMR spectroscopy can be speeded up by the use of simple heteronuclear 2D-experiments.

2.2.6 Quaternary structure determination

Monomeric proteins are characterized by their shape, their surface charges or hydrophilicity/hydrophobicity profile. All these properties and in addition specific chemical functionalities often cause association (quaternary structure) with either identical molecules (homooligomers) or different smaller or larger molecules (ligand-receptor interaction and heterooligomers, respectively). In aqueous solution the association is governed mainly by hydrophobic interaction and charge complementarity which stabilizes the quaternary form.

This multimeric state results in new functional roles or enhanced activities. The quaternary structure is often organized in a symmetrical way to allow the formation of large complexes with only few different monomeric tertiary units, e.g., in the case of the VAT-N protein. VAT-N and other proteins of the aspartic protease family map out an evolutionary path, where duplication and circular permutation of a single copy of the β-barrel VAT-Nn repeat (the N-terminal domain of VAT-N) are responsible to the formation of homodimeric transcription factors and multimeric enzymes which share new functions [115].

The study of the quaternary structure and thus the comprehension of the biological functions motivate from the beginning the purpose of this work. The N-terminal domain of riboflavin synthase (RiSy-N) (see chapter 3) is, indeed, a good candidate for quaternary structure determination with NMR spectroscopy as it forms a homodimer in solution, with riboflavin as a bound ligand. This chapter provides then an overview of some NMR techniques that are commonly used to extract structural information from protein-protein as well as protein-ligand (either an organic or inorganic compound, a cofactor or a metal ion) (Fig. 2.10) or protein-nucleic acid (DNA-RNA) interactions and some of them will be applied to the particular case of RiSy-N in the following chapters.

These techniques are often based on a comparative study between the free state of the molecule and its bound state and they can be separated into two different classes: the first one allows a coarse localization of the interaction site. This includes the use of chemical shift changes to locate ligand binding sites on receptors; of hydrogen exchange rates and paramagnetic agents to map accessible

surface features; of protein dynamics (T_1, T_2 and heteronuclear NOE) to study ligand-induced changes in protein mobility and of anisotropic methods (RDC, CSA) for relative orientation of both molecules. The second class utilizes isotope-edited and isotope-filtered NOESY experiments to probe the interacting sites of the molecules and result, in contrast to the methods cited above, in a precise description of the complex.

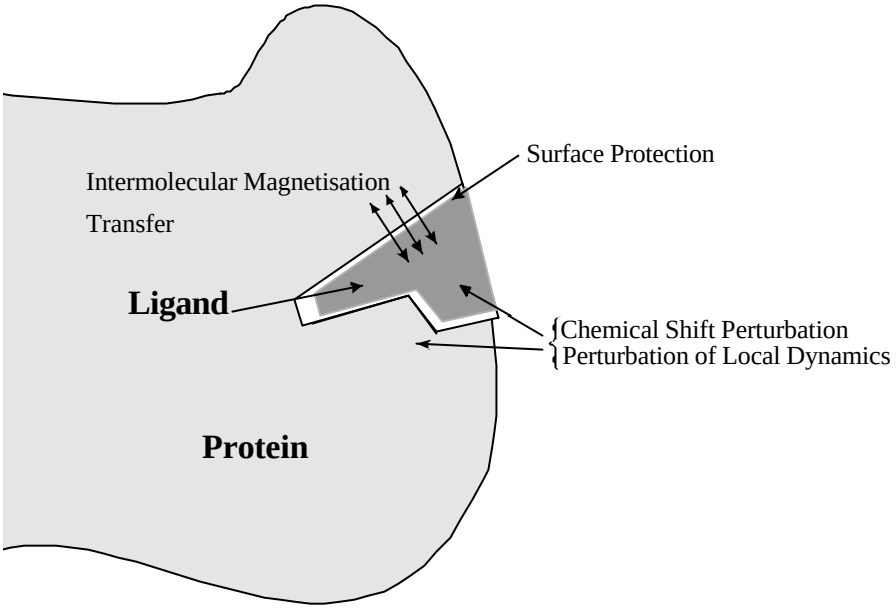


Fig. 2.10: *Schematic representation of some effects due to ligand perturbation. Coarse localization of the interaction site can be achieved via several NMR techniques.*

Chemical shift mapping

An easy and rapid way to observe the interaction between two molecules is to study the chemical shift values. As was mentioned above, the chemical shift is a function of its environment so that it reacts to changes of the local backbone conformations (ϕ, ψ) induced, e.g., by ligand binding. It is also dependent on the changes in anisotropic shielding induced by side-chain rearrangement, H-bonds alteration and others. For the case of a protein with known structure, spectra generally recorded on a ^{15}N-labelled protein sample in presence and absence of a

ligand are compared. Alternatively, a ligand titration allows to follow the chemical shifts changes of the amide groups simply by recording a standard ^{15}N-HSQC [116, 117]. In this case, assuming that the protein and the ligand are in fast exchange (Fig. 2.11), no other experiment is required for the assignment of the concerned shifts that are easy to assign simply following their traces during the titration. In the case of slow exchange, where two different chemical shifts represent the ligand-free and -bound forms, additional experiments are often required for assignment of the residues involved directly or indirectly in the binding site [118-120].

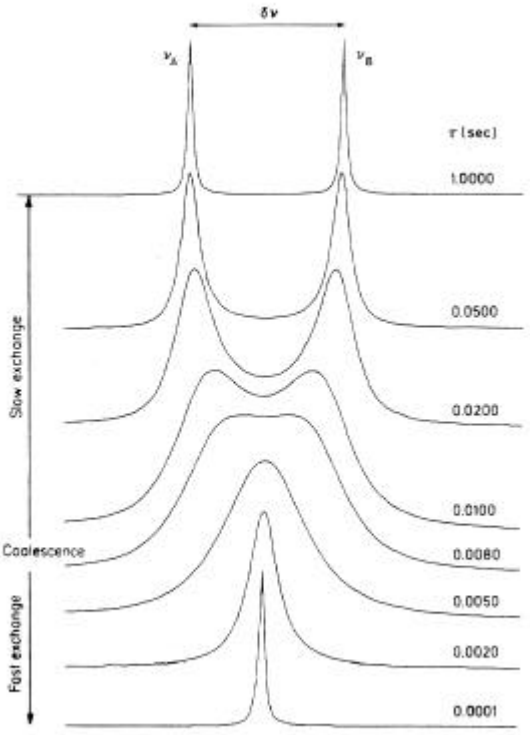


Fig. 2.11: *Typical spectra [121] reflecting chemical exchange at different rates relative to the chemical shifts, for a two-spins system of equally populated sites, e. g., ligand-free and -bound forms.* τ *represents the mean lifetime.*

The binding region can then be mapped onto the structure and represented by, e.g., a color-gradient. This method is intensively exploited in the pharmaceutical

industry for screening random compound libraries for small organic ligands that bind to a target protein [7, 8].

Hydrogen exchange rates (surface mapping)

Solvent accessibility, complexation site and hydrogen bonding can be characterized from hydrogen exchange rate measurements between labile protons (generally, backbone and side-chain amide protons) and the solvent (typically water) [122].

Hydrogen exchange rates can be measured with a MEXICO [123] (Measurement of EXchange rates in Isotopically labeled COmpounds) experiment. After excitation of all protons, ^{13}C- and ^{15}N-^{1}H bound magnetization is filtered out (see below). During the mixing time, only exchange of z magnetization from water to the labile amide protons can occur. The result will be a 2D ^{15}N-HSQC spectrum where only labile proton resonances in exchange with the water are visible. The integration of these correlations reflects directly the hydrogen exchange rates when compared to the reference intensity of the standard HSQC. Experiments for measurement of solvent exchange rates are achieved on the protein in its free and complexed form. Reduced solvent exchange rates take place, then, for the labile protons that are located at the intermolecular interface in the complex, where they are shielded from solvent molecules, compared to their rates in the uncomplexed form. Similarly to chemical shifts mapping, changes in proton exchange rates induced by the presence of the partner molecule can be projected on the protein for coarse localization of the interaction site.

However, this technique can lead to miss-interpretation when applied to the interaction between a small ligand and a macromolecule. Indeed, the protected surface induced by the binding of the ligand is in general very small. It is then difficult to interpret precisely the exchange rates of these few signals. On the opposite, this method is well adapted to the study of the large interaction surface between, e.g., two proteins.

Heteronuclear NOE (steady state NOE)

Structure determination by NMR mainly relies on short range restraints between protons in close spatial proximity (< 5 Å) and torsion angles as described in chapter 2.2.5. In the case of multidomain proteins, protein-ligand interactions, etc., the density of these short distance contacts may be insufficient to characterize the relative positioning of structural elements. This problem can be alleviated by studying protein flexibility. Backbone flexibility measurements allow to evaluate motions and regions of disorder in solution structures, thus providing information on the molecular recognition processes [124].

The flexibility in a protein is measured by heteronuclear-NOE (het-NOE) experiments [125]. Het-NOE takes place when both, the nitrogen and the proton magnetization are along z. A presaturation delay is applied on the amide protons, during which dipolar interactions occur between the saturated amide protons and their bound nitrogen. It follows a ^{15}N-HSQC where the intensity of the NH peaks is directly correlated to the protein flexibility.

Practically, motional parameters are determined with and without the presence of a ligand, enabling the recognition of the binding site as well as the regions that exhibit conformational changes due to ligand binding [126], simply by data comparison. In case of the study of multidomain proteins (where no previous comparison is possible), parts of the protein that do not present a secondary structure, e.g., coils, should have an undefined structure. The involvement of these regions in a site of dimerisation can be thus easily observed. Indeed, high het-NOE rates that reflect a low flexibility are expected.

Isotope filter experiments

The more precise way to study the mode of binding between two molecules is still based on the NOE effect. As a function of the kind of molecular interactions, the method has to be able to study a particular biological system easily. However, such a study is difficult and sometimes very expensive. Indeed, the analyse of the

NOEs are only possible if the two molecules can be distinguished on the basis of differential labeling as shown in Fig. 2.12.

A differential labeling of the interacting molecules allows a discrimination of the protons on the basis of the heavy atoms to which they are bound, and thus the use of isotope filter experiments [127] to separate before and after the NOE mixing time the two different types of protons.

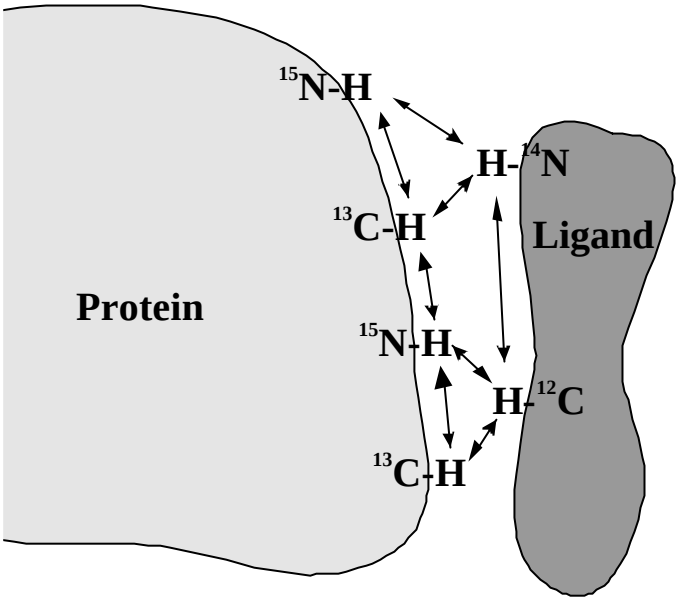


Fig. 2.12: ***Scheme of isotope-filtered and -edited heteronuclear NMR as applied to a macromolecular complex. In this example, the protein is doubled labeled while the ligand is at natural isotopic abundance.***

There are many possibilities to construct an isotope filter experiment, i.e., the dimensionality of the spectra, the nuclei that will be recorded in each dimension and the number of filter sequences to apply in the dimensions have to be chosen. As a definition, the mention "^{13}C-filter" means that the ^{1}H signals bound to ^{13}C are removed; a "^{13}C-edited" experiment in contrast records exclusively these protons, and an experiment where the heteronuclei itself are detected will be called "^{13}C-separated".

Filter experiments based on a difference technique

The method differentiates ^{1}H attached to an isotopically enriched heteronucleus from those attached to a heteronucleus with a natural abundance isotope. Indeed, the ^{1}H-^{15}N and ^{1}H-^{13}C spin pairs (represented by I-S in the product operator formalism [128]) evolve under the heteronuclear coupling and under the heteronuclear chemical shift after a radio frequency pulse is applied to the heteronuclei. On the opposite, the Hamiltonian of ^{1}H-^{14}N and ^{1}H-^{12}C pairs (represented by H in the product operator formalism) doesn't evolve under these effects. A simple implementation of a heteronuclear 180° pulse allows to discriminate between the coherence of these two ^{1}H-spins:

$$2I_xS_z \xrightarrow{180°_x(X)} -2I_xS_z$$

$$H_y \xrightarrow{180°_x(X)} H_y$$

Two experiments are then alternatively applied; one with the 180° pulse and one without. A linear combination of these two data sets [129] discriminates then between the information based on a filtered or edited experiment. As an advantage, both information are present, when stored separately, due to the fact that no magnetization is destroyed. The inconvenience of such filters, called X-filters, results in the absence of a tunable delay that limits their application to systems with homogenous heteronuclear coupling constants. Also, the quality of the spectra is very dependent on those of the pulse calibrations that are entirely responsible for the quality of the filter.

The X-half-filter can be in contrast tuned via adjustment of the length of an evolution delay, for a specific magnitude of an I-S coupling constant. This delay is however tuned to a particular spin system so that its efficiency is a function of the homogeneity of the heteronuclear coupling constant. The X-half-filter (Fig. 2.13) [130] and corresponding X-double-half-filter (applied in both dimensions of a 2D spectrum), show generally a good performance when applied to ^{1}H-^{15}N spin pairs.

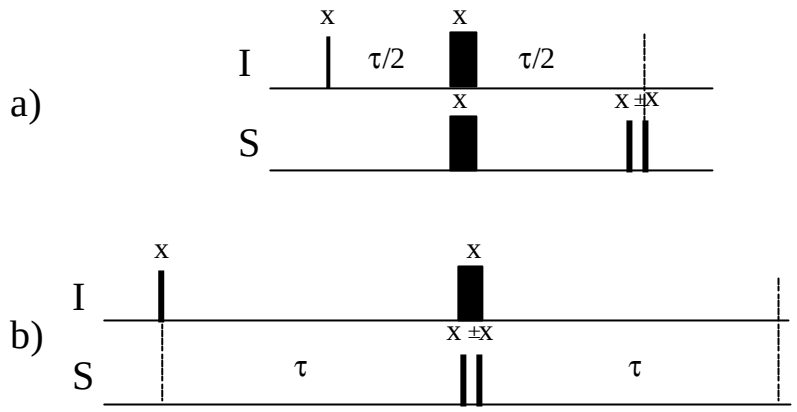


Fig. 2.13: *Two X-half-filters: the delay* ***t*** *is set to 1/(2*$^1J_{IS}$*): a) basic scheme, b) modified scheme that refocuses anti-phase coherence.*

As shown for the X-filter, only the sign of the anti-phase coherence is affected by the two successive 90° pulses. Due to the phase alternation of the second 90° pulse, from transient to transient, it is possible to simulate the presence or absence of the 180° pulse explained for the X-filter. Indeed, the effective flip angle is alternatively set to 180° or 0°. The disadvantage of these filters is the intensity loss due to the implantation of an additional delay in the sequence, especially in case b) of Fig. 2.13. This filter requires a length twice as in case a) to refocus the anti-phase coherence into an in-phase one, allowing thus a X-nucleus decoupling. However, the main difficulty in this kind of filter is to set the delay properly for filtration of a large background of protons from labeled macromolecules in a complex, in order to observe a smaller number of ^{1}H correlations from the unlabelled ligand partner.

Isotope filters based on purging schemes

The isotope filters based on purging schemes (low-pass J-filters [131]) allow only the collection of heteronuclear filtered data. The principle is to destroy the heteronuclear coherence sending it into multiple quantum coherence order. They are also easy to implement and the filtration is achieved in a single transient and does not rely on a difference technique thus resulting in high spectral quality: better filtration and less artifacts. The sequence is based on either an HMQC- or

INEPT-type transfer where the last ^{1}H pulse is omitted to fix the coherence on double quantum coherence. These filters have the advantage of obtaining the same result as the X-half-filter for a filtration delay that is about 50 % shorter. It allows then for a similar sensitivity the concatenation of two successive filters where two filtration delays can be set for the suppression of spin pairs that evolve under different coupling constants. Thus, for the first time it is possible to destroy with good efficiency the magnetization coming from the ^{1}H-^{13}C spin systems. The double tuned filters introduced by Gemmecker *et al.* (Fig. 2.14) [132] incorporate also a ^{1}H spin-lock purge pulse to dephase through radio frequency inhomogeneity any remaining anti-phase coherence that escaped conversion into heteronuclear multiple quatum coherence (MQC) by the ^{13}C pulse.

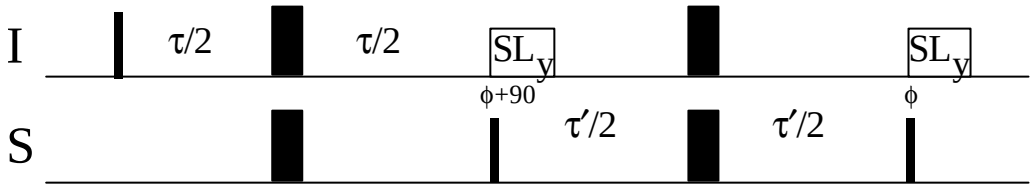


Fig. 2.14: *"Doubly tuned" filter motive [132]. The filter delays* ***t*** *and* ***t'*** *are set to two different values in respect with the size of two different heteronuclear coupling constants. The efficiency of this filter is improved by the use of spin-lock purge pulses.*

Later, similar techniques used purged filters based on a PFG z-filter [133]. Magnetization of protons attached to X-nuclei are converted to anti-phase coherence via an INEPT-like pulse sequence, while the magnetization of protons not bound to a heteronucleus stays at H_y (Fig. 2.15). The second 90° ^{1}H pulse sends the H_y magnetization back to z while leaving unaffected the anti-phase term in the x-y plane. The pulsed field gradient dephases then the transversal magnetization without affecting the longitudinal magnetization that can be detected in the following.

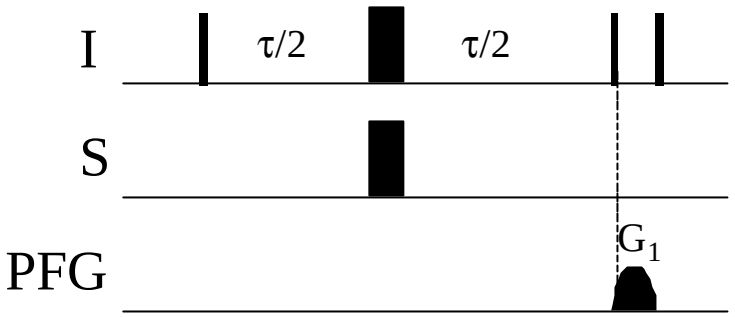


Fig. 2.15: *PFG z-filter pulse sequence: the delay t/2 is set to 1/(4$^1J_{IS}$). All pulses are applied along the x-axis. The PFG G_1 purges all transverse magnetization.*

Also, it is possible to use sequences as CT, SHIT and others for concatenation of filter schemes and incrementation times in order to minimize intensity loss due to addition of several filter modules that are present in one, two or more dimensions. More uniform inversion performance was achieved using composite pulses such as $90°_x$-$180°_y$-$90°_x$ compensating for the miss-calibration of the X-pulses. Pulse sequences such as wide-band inversion [134] or band-selective inversion pulses [135] were developed to filter simultaneously aliphatic and aromatic carbon bound protons. Finally, adiabatic pulses were proposed to solve the problems of filter band width. The adiabatic pulses are frequency swept and thus increase the quality of inversion pulses for the ^{1}H-^{13}C spin pair, due to a quite linear correlation between the size of $^1J_{HC}$ and the carbon chemical shift [136].

In conclusion, description of the binding mode of two or more molecules is very important for the understanding of biological functions. It is a challenge that can be solved by NMR spectroscopy. A very large set of NMR techniques allows the structure determination of very different biological complexes, and recent programs dedicated to structure calculation offer the possibility to discriminate intra- from intermolecular NOE correlations [137, 138]. The limiting factor for the application of these methods can be the cost and difficulty of NMR sample preparation. Differential labeling is often required. Indeed, the investigation of protein-ligand interactions requires prior assignment of the ligand resonances. This can be performed on unlabeled protein with a ^{13}C enriched ligand when the isotope labeling of the ligand is possible. This sample can also be used for discrimination

between spins of both molecules and to solve sometimes difficulties in filtration of protein ^{13}C magnetization. Additionally, the same information can be achieved with a double labeled protein and unlabeled ligand. The study of multimeric proteins is more complicated due principally to the symmetry properties of the complex. For example, in case of a homodimer, expression of a cross-labeled (or orthogonally labeled) protein allows the use of well-known edited NOESY experiments. The preparation of such a sample consists generally in the separate expression of single ^{13}C- and ^{15}N-labeled homodimers. Then, heating denaturation permits a statistical recombination of the monomers in the proportion 1:2:1; i.e., 25% each of the singly ^{13}C-labeled dimer and its ^{15}N homologue, and 50% of the mixed ^{13}C-^{15}N complex. Such a sample is however not easy to produce at NMR concentrations, because the concentration of the mixed dimer will dramatically decrease if the concentration of both singly labeled dimers is different (before the denaturation step), resulting in unobservably weak NOE cross-peaks.

3 Structure Determination of the N-Terminal Domain of Riboflavin Synthase

3.1 Biochemical Background

The pathway for biosynthesis of riboflavin (Fig. 3.2) has attracted considerable attention over many years. This is certainly mainly due to the difficulties in understanding an "unusual" reaction mechanism, with many attempts to explain it, giving rise to dozens of publications.

Riboflavin (or vitamin B2) was long confused with vitamin B1 (thiamine). First in 1920, A. D. Hemmett remarked the presence of another active substrate after heating of a yeast and destruction of the thiamine [139]. Later, in 1932, Warburg *et al*. isolate the riboflavin, which was called "yellow enzyme" due to its color (from Latin, flavus = yellow) [140]. It is also found in several natural substances (actually, riboflavin is commercially used as dye in food chemistry). In 1933, R. Kuhn could isolate the "yellow enzyme" from milk [141]. He then synthesized the riboflavin in 1935 and solved its structure ($C_{17}H_{20}N_4O_6$) (Fig. 3.1) [142]. Riboflavin is now known to be synthesized by a large number of plants and microorganisms. In 1945, Peterson [143] listed some 75 bacterial species that were observed to produce this vitamin. The term "riboflavin" was definitively adopted in 1952 by the nomenclature commission of biochemistry (Riboflavin stands for ribose + flavin). This water soluble vitamin plays an important role in electron transfer processes after transformation into flavin mononucleotide, FMN, and flavin-adenine dinucleotide, FAD (both coenzymes of several flavoproteins) as represented:

$$\text{Riboflavin} \underset{\text{Phosphatase}}{\overset{\text{Flavokinase + ATP}}{\rightleftharpoons}} \text{FMN} \underset{\text{Pyrophosphatase}}{\overset{\text{Pyrophosphorylase}}{\rightleftharpoons}} \text{FAD}$$

Riboflavin is an isoalloxazine derivative, substituted at C_5 by the sugar ribitol. FMN and FAD are directly derived from riboflavin after phosphorylation of the

ribitol. Thus, riboflavin is a precursor of essential redox cofactors in all cellular organisms and can be found in different oxidative forms as represented in Fig. 3.1.

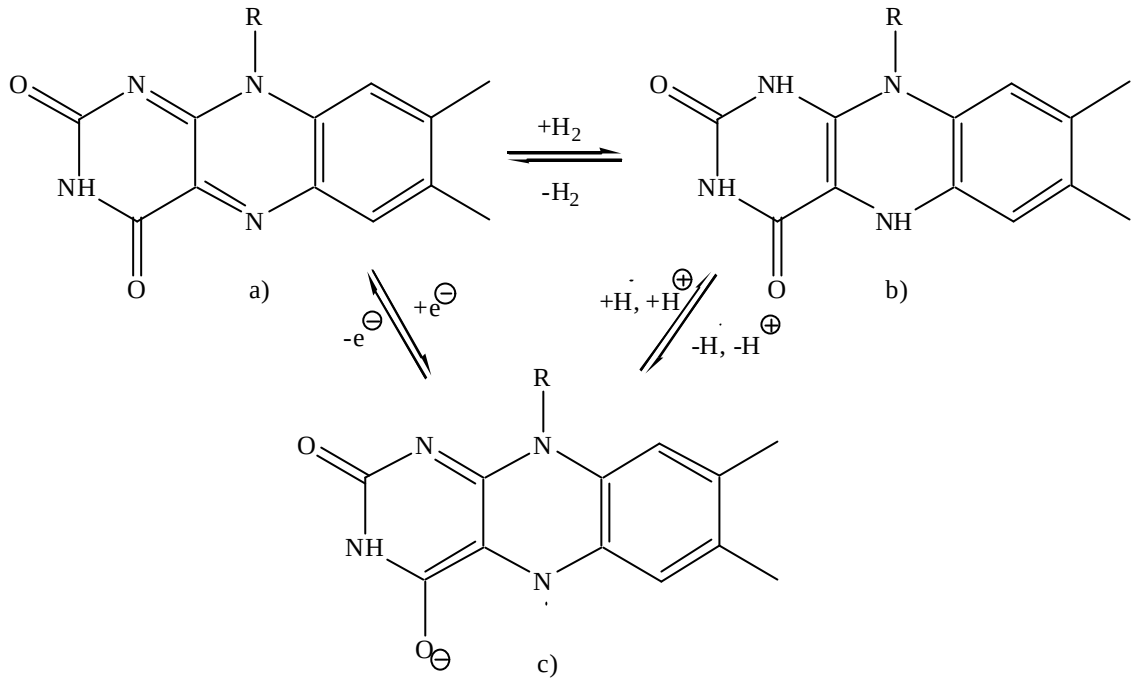


Fig. 3.1: *Principal oxidative forms of riboflavin. a) oxidized form, b) reduced form, c) semiquinone form.*

Riboflavin synthase belongs to a protein family together with lumazine binding antenna proteins and yellow fluorescent protein (YFP). Both serve as phototransducers in bacterial bioluminescence [144] [145] and show a high degree of sequence homology with riboflavin synthase.

In the evolution of the species, several animals lost the ability to biosynthesize riboflavin. Human beings generally don't have any vitamin B2 deficiency because of a large presence of riboflavin in food (beer yeast and others, cheese, egg, some fruits and vegetables, milk, fish…). However, Gram-negative bacteria and certain yeasts are virtually unable to absorb riboflavin and riboflavin derivatives from their environment and are therefore absolutely dependent on the endogenous biosynthesis of the vitamin. The enzymes of the riboflavin pathway can hence be considered as potential targets for the development of anti-bacterial and anti-fungal agents.

The pathway of riboflavin biosynthesis in *Escherichia coli* is summarized in Fig. 3.2. Riboflavin synthase catalyses the final step in the biosynthesis of riboflavin which involves the dismutation of 6,7-dimethyl-8-ribityllumazine (lumazine, **1**) (for a review see Bacher *et al.* [146]). The reaction involves the transfer of a C_4 unit between two identical substrate molecules [147]. The xylene ring of riboflavin is obtained by head-to-tail assembly of two C_4 units provided by the donor and acceptor substrate molecules [148-150]. The second product of the reaction, 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione (**3**), can be recycled in the biosynthetic pathway via lumazine synthase [147].

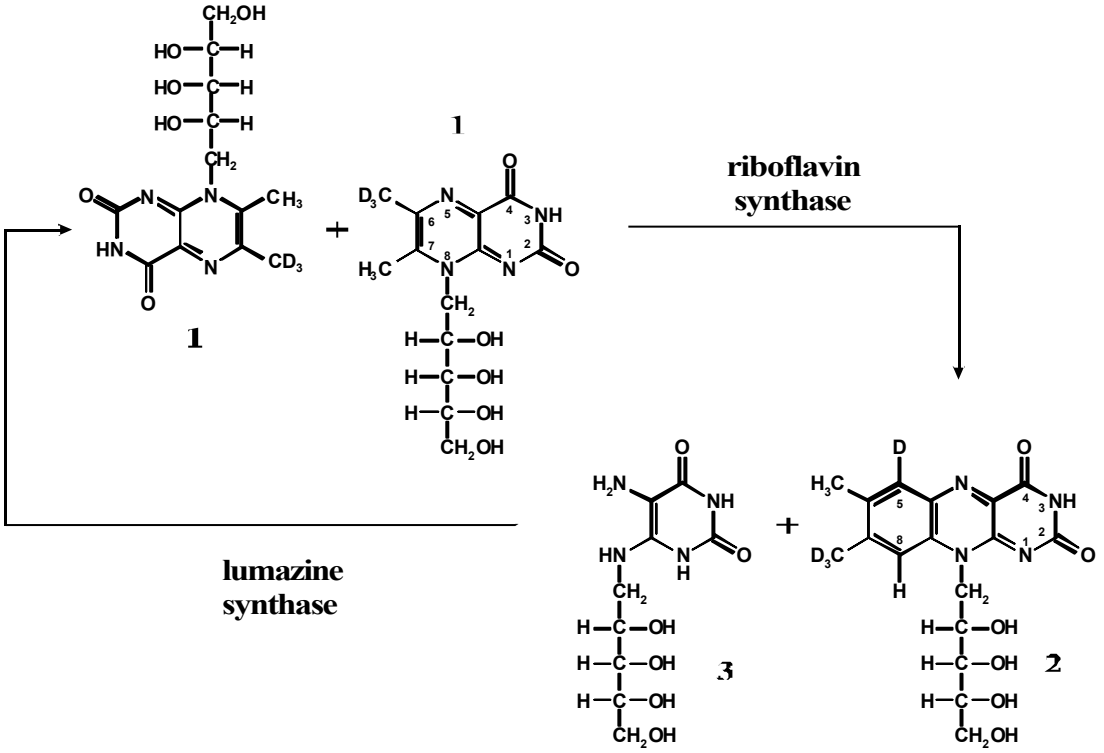


Fig. 3.2: *The bacterial biosynthesis of riboflavin. Two molecules of 6,7-dimethyl-8-ribityllumazine (lumazine; 1) form riboflavin (2) by a dismutation reaction. 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione (3) is a recyclable side product. The regioselectivity of the dismutation reaction has been determined by deuterium labeling as indicated, with two deuterium atoms being transferred to the solvent* [148].

Riboflavin synthase of *Bacillus subtilis* and *Escherichia coli* present remarkable differences. *B. subtilis* has two different riboflavin synthases characterized by the

subunit structures α_3 (light enzyme) and $\alpha_3\beta_{60}$ (heavy enzyme). The light enzyme is a trimer of identical α subunits, whereas the heavy enzyme consists of an α subunit trimer enclosed in a lumazine synthase capsid with icosahedral symmetry which is composed of 60 β subunits [151, 152]. In contrast, lumazine synthase of *E. coli* does not form a complex with riboflavin synthase or any other protein of the riboflavin pathway and is therefore an empty icosahedral capsid. Whereas the heavy enzyme is known in considerable details [153-155], little is known about the structure of the 23 kDa α subunit.

The riboflavin synthases of *B. subtilis* and *E. coli* are homotrimeric proteins [10, 156] with a total molecular mass of approximately 75 kDa. The monomeric unit is characterized by internal sequence similarity; for example, the N- and C-terminal domains of the enzyme from *Escherichia coli* are 26 % identical (Fig. 3.3) [157-159]. This strongly suggests the presence of two topologically similar folding domains [10]. In line with this hypothesis, low-resolution crystallographic analysis suggested a pseudo-D_3 symmetry for the homotrimeric protein, indicating that riboflavin synthase (RiSy) behaves as a pseudo dimer [160].

The Bacher group has recently found that the N-terminal domain of *E. coli* riboflavin synthase, comprising amino acid residues 1-97 (RiSy-N), forms a soluble, stable protein when expressed in a recombinant *E. coli* strain. Moreover, the N-terminal domain forms a homodimer in solution which shows native-like substrate binding [161]. In this thesis, the solution structure of this homodimer is presented in the presence of riboflavin as a bound ligand, as determined by multidimensional, heteronuclear NMR spectroscopy. Several pieces of evidence suggest that the RiSy-N homodimer forms a good structural model for the RiSy monomer. The structure therefore provides insights into both the mechanism of riboflavin synthesis and the structure of the functional RiSy trimer.

Fig. 3.3: *An alignment of riboflavin synthase sequences* (RISA). *Residues conserved or homologous in both RiSy-N and RiSy-C are shown in red and blue, respectively, while residues highly conserved within one domain are shown in purple. Also shown are the sequences of three lumazine binding antenna proteins* (LUXP/LUXY) *from marine bacteria which are non-catalytic monomers which bind one substrate molecule. The alignment shows the similarity of the N- and C- domains. Areas coloured green and yellow show residues expected to participate in the binding sites of RiSy-N and RiSy-C, respectively. Organism codes are:* ACTPL, *Actinobacillus pleuropneumoniae;* BACSU, *Bacillus subtilis;* ECOLI, *E. coli (from which the residue numbering is derived);* HAEIN, *Haemophilus influenzae;* PHOLE, *Photobacterium leiognathi;* PHOPO, *Photobacterium phosphoreum;* VIBFI, *Vibrio fischeri.*

3.2 Materials and Methods

3.2.1 Sample preparation

The expression and purification of the recombinant N-terminal domain of riboflavin synthase (amino acid residues 1-97) of *E. coli* was performed at the Technische Universität München in the group of Prof. Dr. Dr. A. Bacher by Dr. Sabine Eberhardt and Dr. Holger Lüttgen [161]. The purified protein was dialysed extensively against 50 mM phosphate buffer, pH 7.3, containing 1 mM riboflavin. Additional riboflavin was added as a solid in order to saturate samples with ligand. For the NMR structure determination of RiSy-N, [U-^{13}C, ^{15}N]- and [U-^{15}N]-labelled protein samples with unlabelled riboflavin as a bound ligand were used with a typical sample concentration of 1 mM and in addition ca. 10 % D_2O (appendix B2).

3.2.2 NMR Spectroscopy and structure calculations

All spectra were recorded at 300 K at 600 or 750 MHz on Bruker DMX600 and DMX750 spectrometers, respectively. Other spectrometer characteristics as well as calibrated spectral widths used for RiSy-N are reported in appendix A1 and appendix B1, respectively. The spectra were processed with XwinNMR ['X-WINNMR V3.0', Bruker, Karlsruhe] on an Octane R10000 Silicon Graphic (SGI) work station with a CPU rate of 175 MHz. Aurelia ['AURELIA V2.8.11', Bruker, Karlsruhe] and Pasta ['PASTA V2.0', Technical University of Munich] [66] were extensively used for the chemical shift assignment. Structures were calculated in XPLOR ['XPLOR V3.8.51', Yale University] using generally standard simulated annealing protocols [162], and visualized in Insight II ['Insight II V2000', Accelrys, Princeton]. The punctual use of other programs will be appropriatly mentioned. Pulse programs and parameter conditions are summarized in appendix B3 and appendix C1.

3.2.3 Database details

The final structure set and the regularised average structure for the RiSy-N homodimer have been deposited in the RCSB PDB under the codes 1I18 and 1HZE, respectively. A full list of chemical shifts has been deposited in the BioMagResBank (BMRB) database under the code 4954.

3.3 Chemical Shift Assignments

3.3.1 Oligomerization state of RiSy-N

The apo-form of the N-terminal domain of RiSy has been shown by sedimentation studies to form a stable homodimer in solution [161]. To confirm the oligomeric purity of the samples under the conditions of the NMR experiments the molecular diffusion coefficient for the protein was measured by diffusion experiments.

Diffusion data provide details on molecular organization, and reflect the quality of the sample through its oligomarization state (aggregation, fragmentation for example) [163]. Diffusion rates are sensitive to structural changes as well as to dissociation/association phenomenon. Diffusion experiments can be performed and fully interpreted before the laborious assignment process.

The basic principles of diffusion measurement are simple: after an initial 90° proton pulse, spins at different precession rates dephase in the x,y-plane during the diffusion delay. A 180° pulse inverts the magnetization which will be refocused after application of an identical diffusion time. A spin echo (SE) is thus obtained. The 90°-180° echo sequence was introduced by Carr and Purcell [164] and later modified into the Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence [165]. This sequence is also well known for measurements of transverse relaxation rates by NMR. However, more advanced pulse sequences were developed for several reasons. First, diffusion phenomena interact with the relaxation and it is difficult to separate diffusion contributions to spin relaxation from other relaxation pathways. An increase of the diffusion delay results in a large contribution of undesired transversal relaxation effect (in particular for large biological compounds), thus leading to non-interpretable diffusion rates. Then, additional energy from punctual heating of the sample, due to spectrometer instabilities or application of pulse field gradients, increases the internal kinetic energy and thus the molecular motions. It leads to inhomogeneous dephasing of the coherence during the spin diffusion

delay, which is directly observable on the diffusion coefficient values. Fig. 3.4 displays different pulse sequences for the measurement of diffusion rates.

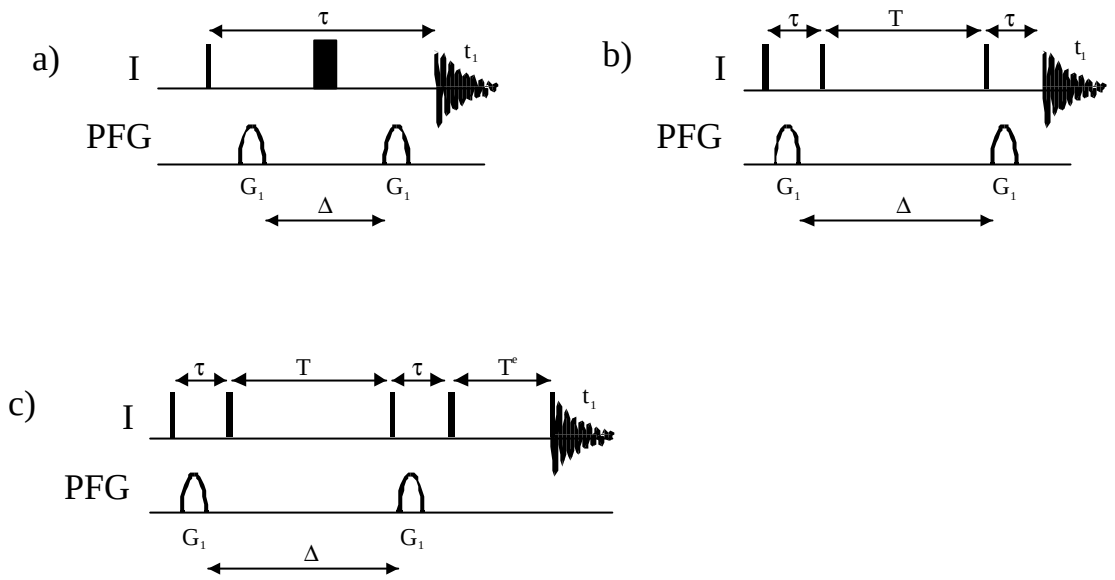


Fig. 3.4: *Pulse sequences using pulsed field gradients for the measurement of diffusion. a) the Pulsed field Gradient Spin Echo technique (PGSE). b) the STimulated Echo method (STE-PFG). c) The Longitudinal Eddy-current Delay sequence (LED-PFG).*

The PGSE experiment [166] was the first experiment which used pulsed field gradients for diffusion measurement (Fig. 3.4a). It resembles the standard SE-pulse sequence, but allows the application of a diffusion period without incrementation of any delay. Indeed, coherence dephasing can occur either during a delay (that could also be the gradient length) or, for suppression of transversal relaxation, by increasing the gradient strength as represented by the following equation:

$$\Delta\phi = \gamma G(\Delta z)t \qquad \textbf{eq. 3.1}$$

where $\Delta\phi$ represents the spin phase changes, γ the gyromagnetic ratio of protons, G the linear field gradient applied in the z-direction and Δz the molecular displacement achieved during the delay t. The STE-PFG experiment [167] is more sensitive. Indeed, for large proteins, T_2 is generally much shorter than T_1. Also, in order to diminish the intensity loss due to extensive relaxation, the delay during

which the coherence is in the transversal plane is reduced to the minimum (2τ) (Fig. 3.4b). Thus, the coherence is dephased by the gradient in the transversal plane (phase labelling), send back on z during the diffusion period T, and refocused in the transversal plane by the second gradient.

Finally, Gibbs and co-workers [168] proposed in 1991, with the LED-PFG pulse sequence, to implement an extra gradient recovery delay T_e that lets enough time to the eddy current fields to decay to an acceptable level for undistorted signal acquisition (Fig. 3.4c).

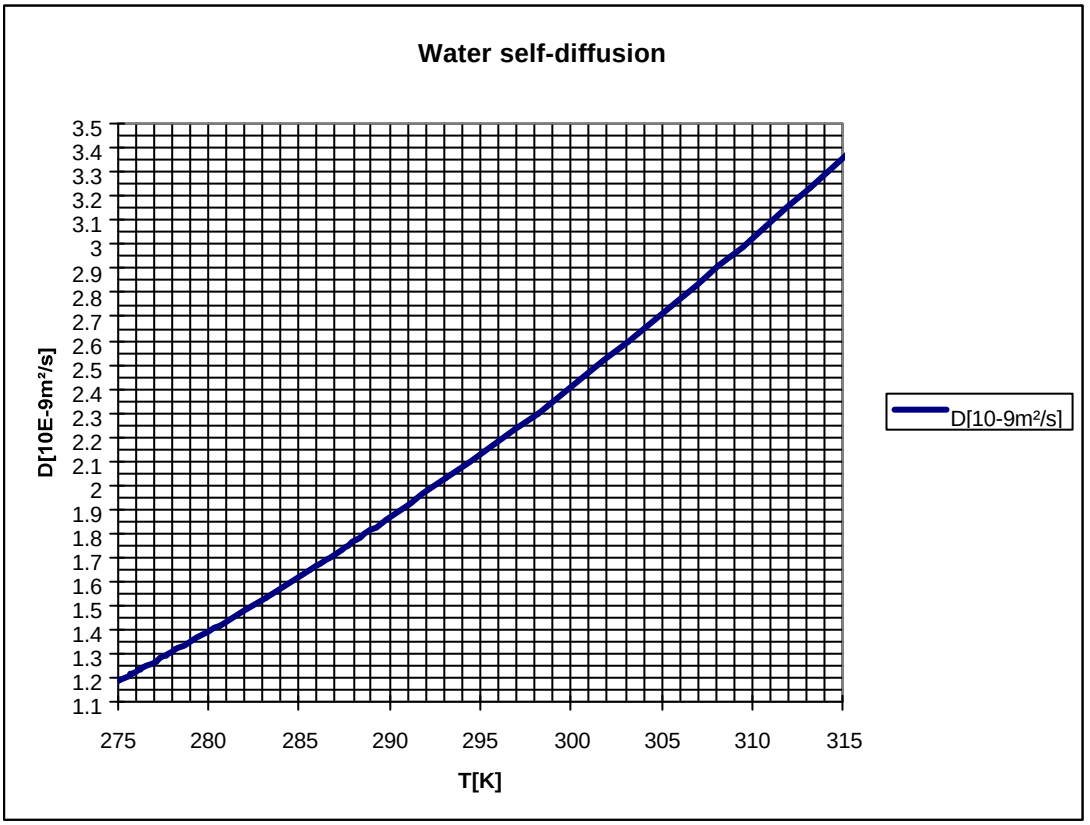


Fig. 3.5: *Self-diffusion rates calibrated for the water as a function of the temperature.*

The pulse program applied for diffusion measurements on RiSy-N employs the latter LED-PFG sequence and a combination of z-spoil gradients (during T- and T_e-periods) together with an optimised phase cycling for a better artefact suppression (see appendix C1). Two pseudo-2D diffusion experiments were obtained from a unique [U-^{15}N]-labelled protein with unlabeled riboflavin. The

first experiment measures the water self-diffusion. Water diffusion rates will be compared to calibrated values (Fig. 3.5) and serve as reference before measurements of the protein diffusion. Both spectra are recorded under similar conditions with 128 experimental points representing the diffusion curves at 300K.

Analysis of the diffusion curves was done as described by Stejskal and Tanner [166], in the approximation of free molecular self-diffusion in an isotropic solvent. The signal intensity, in case of a PFG-NMR experiment is:

$$I(t) = I_0 \exp[-D \cdot x] \text{ with } x = \gamma^2 \cdot G^2 \cdot \delta^2 \cdot (\Delta - \delta/3) \qquad \textbf{eq. 3.2}$$

where D is the diffusion coefficient; G and δ the strength and the length of the diffusion gradients, respectively. As described above, the diffusion is achieved through a gradient ramp. G takes 128 different values between 2 and 40 % of the gradient amplifier strength for the water and 2 and 80 % for the protein; δ = 2ms. Δ is the time between the diffusion gradients and thus represents the diffusion delay, it is equal to 28 ms respectively 128.7 ms for the water and the protein. Finally, γ is the effective ^{1}H gyromagnetic ratio:

$$\gamma = 4258 \times g_{scale} \times g_{shape} \text{ (in Hz/cm).} \qquad \textbf{eq. 3.3}$$

g_{scale} is the calibrated gradient strength: 0.67 G/cm for the TXI probe head on the DMX-601 spectrometer; g_{shape} is equal to 1 for square gradient shapes and, in our case, $2/\pi$ for sine gradient shapes. Other parameters were set to standard values and are thus not mentioned here. Diffusion curves were fitted according to eq. 3.2 using the T_1/T_2 relaxation analysis module contained in XwinNMR and shown in Fig. 3.6.

Self-diffusion of the water is measured to be 2.29 ($\pm$ 0.10) $\times 10^{-9}$ m^2s^{-1} that matches perfectly the theoretical value within the experimental error (2.4×10^{-9} m^2s^{-1}, Fig. 3.5) and thus reflects the quality and stability of all spectrometer parameters. The value obtained of 1.05 ($\pm$ 0.05) $\times 10^{-10}$ m^2s^{-1} at

300 K for RiSy-N can be compared to other proteins. The solution structure of HNL (ca. 20 kDa) [169] and VAT-N (20.6 kDa) [170] was solved in our laboratory using NMR spectroscopy. Diffusion rates show 1.05 (± 0.05) × 10^{-10} m^2s^{-1} at 297.5 K for HNL, and 1.55 × 10^{-10} m^2s^{-1} for VAT-N at 320 K. In reference to these values, the diffusion coefficient of RiSy is in perfect agreement with the molecular mass of the homodimer (20 kDa) and excludes significant exchange between higher or lower oligomeric forms. In line with this a single set of backbone resonances was observed for the domain. Moreover, given that the samples contained riboflavin in excess, the single set of signals also indicates that this dimer contains two symmetrical binding sites.

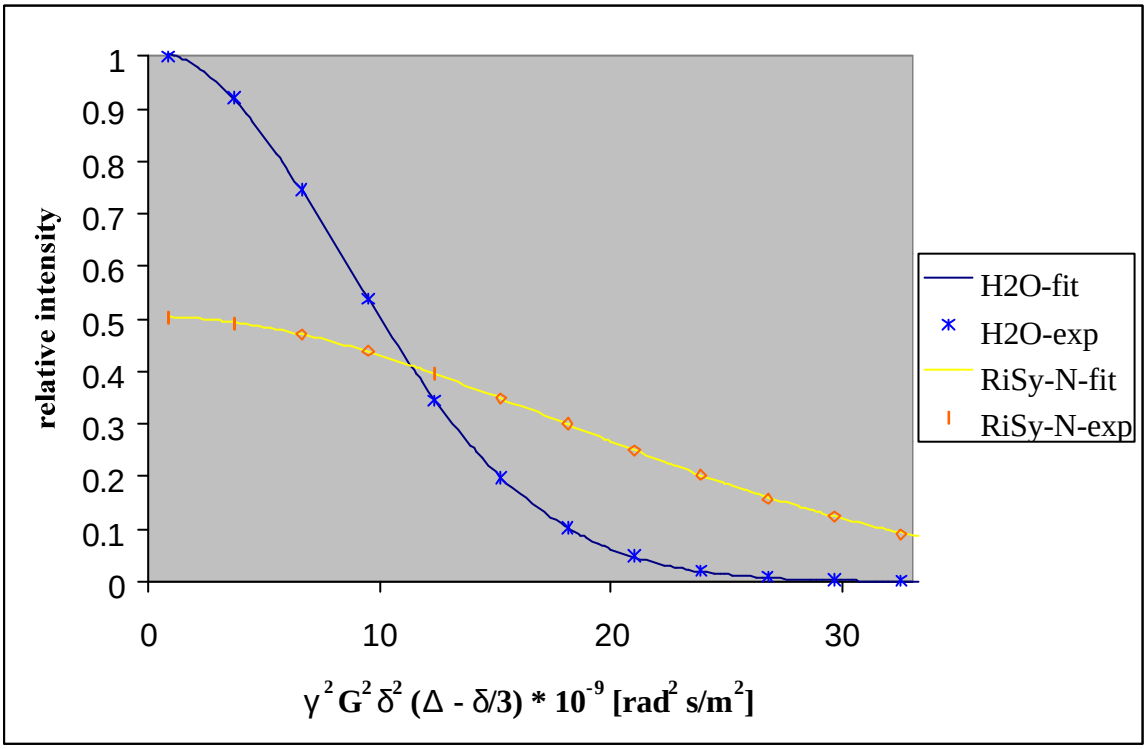


Fig. 3.6: *Relative intensity of self-diffusion curves measured on a [U-^{15}N]-RiSy-N labeled sample. Both curves are shown as a function of* $\gamma^2 G^2 \delta^2 (\Delta - \delta/3) \times 10^{-9}$. *Only few experimental points are shown for sake of clarity.*

A series of riboflavin titration HSQC experiments were also acquired and close examination showed no traces of signals due to unliganded RiSy-N in the samples with ligand in excess. Thus, both binding sites of the dimer are predominantly

occupied. In contrast, samples with RiSy-N in excess showed multiple signal sets, indicating slow exchange between apo- and holo-forms of the protein. It was also possible to assign the backbone amide resonances of the apo-form of the protein via comparison of HNHA spectra acquired on liganded and unliganded ^{15}N-labelled samples. Perturbations of amide ^{1}H and ^{15}N chemical shifts upon ligand binding are discussed in chapter 3.6.

3.3.2 Resonance assignment

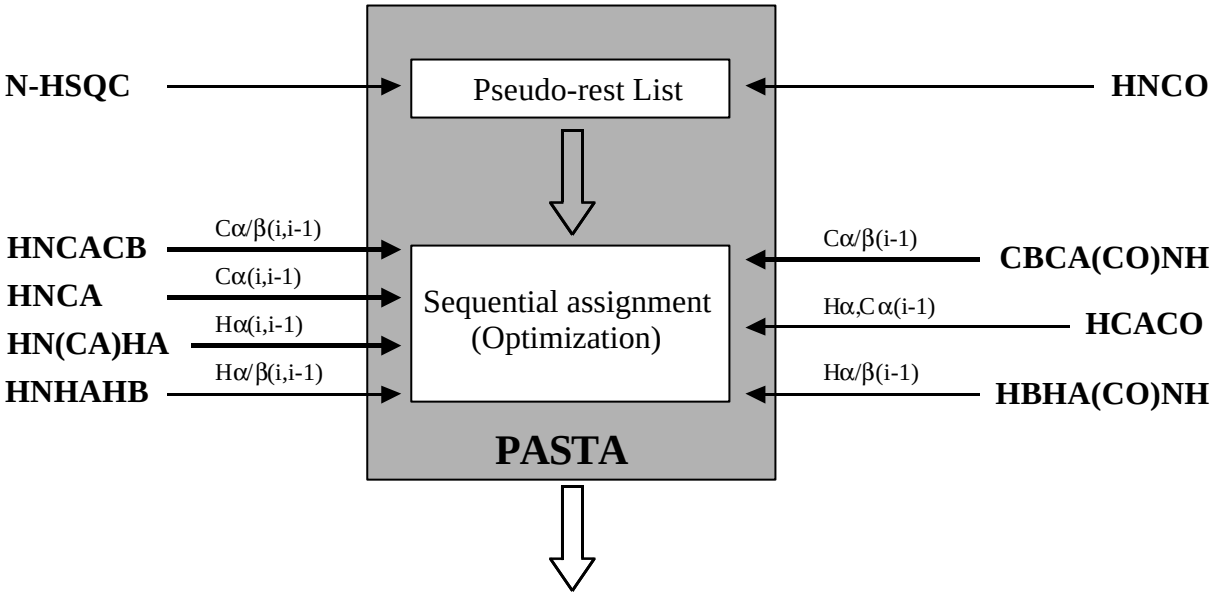
Backbone resonances of riboflavin liganded RiSy-N were assigned with the automatic assignment program PASTA [66] using sequential C_α, C_β, C', H_α and H_β chemical shift information derived from an array of triple-resonance experiments (HNCA, HCACO, HNCACB, CBCA(CO)NH, HN(CA)HA, HNHAHB and HBHA(CO)NH), as represented by Fig. 3.7. As was discussed in chapters 2.2.2 and 2.2.3, the assignment of β-spin resonances is the key of the assignment process allowing to rely the spin systems of the side-chains to their respective backbones.

Additionally, carbonyl carbon assignments were available from an HNCO and HCACO experiments. A HNHA experiment performed on the ^{15}N-labelled sample was used to measure $^{3}J_{HNH\alpha}$ coupling-constants.

Backbone resonances for all residues were assigned, with the exception of the C´ resonances of K18 and L28, which precede prolines, and M1, and the completely assigned ^{15}N-HSQC of liganded RiSy-N is shown in Fig. 3.8.

Assignment of side-chain ^{1}H and ^{13}C resonances was completed using a strategy where C_β and H_β assignments obtained during sequential assignment served as starting points for analysis using a combination of 3D H(C)CH-TOCSY and 3D (H)CCH-COSY experiments (see chapter 2.2.3). A small number of side-chain protons, generally the amino groups of arginine and lysine residues remained unassigned due to weak signals or ambiguity. Similarly, aromatic resonances were assigned by NOE contacts and a small number of these remain unassigned.

Stereospecific assignment, made by consideration of NOESY cross-peak patterns (Fig. 2.8), was possible for the prochiral H_β protons of 41 out of 71 residues and for the prochiral methyl groups of all valine residues. A Table containing the 99 % of backbone and 94 % of side-chain shifts assigned for RiSy-N has been deposited in the BMRB database (accession number 4954).



After manual verification of the automatic assignment
with the help of NOE connectivities:

96 (of 97) assigned residues

Fig. 3.7: *Schematic procedure of the sequential assignment of RiSy-N with the help of the assignment program PASTA* [66].

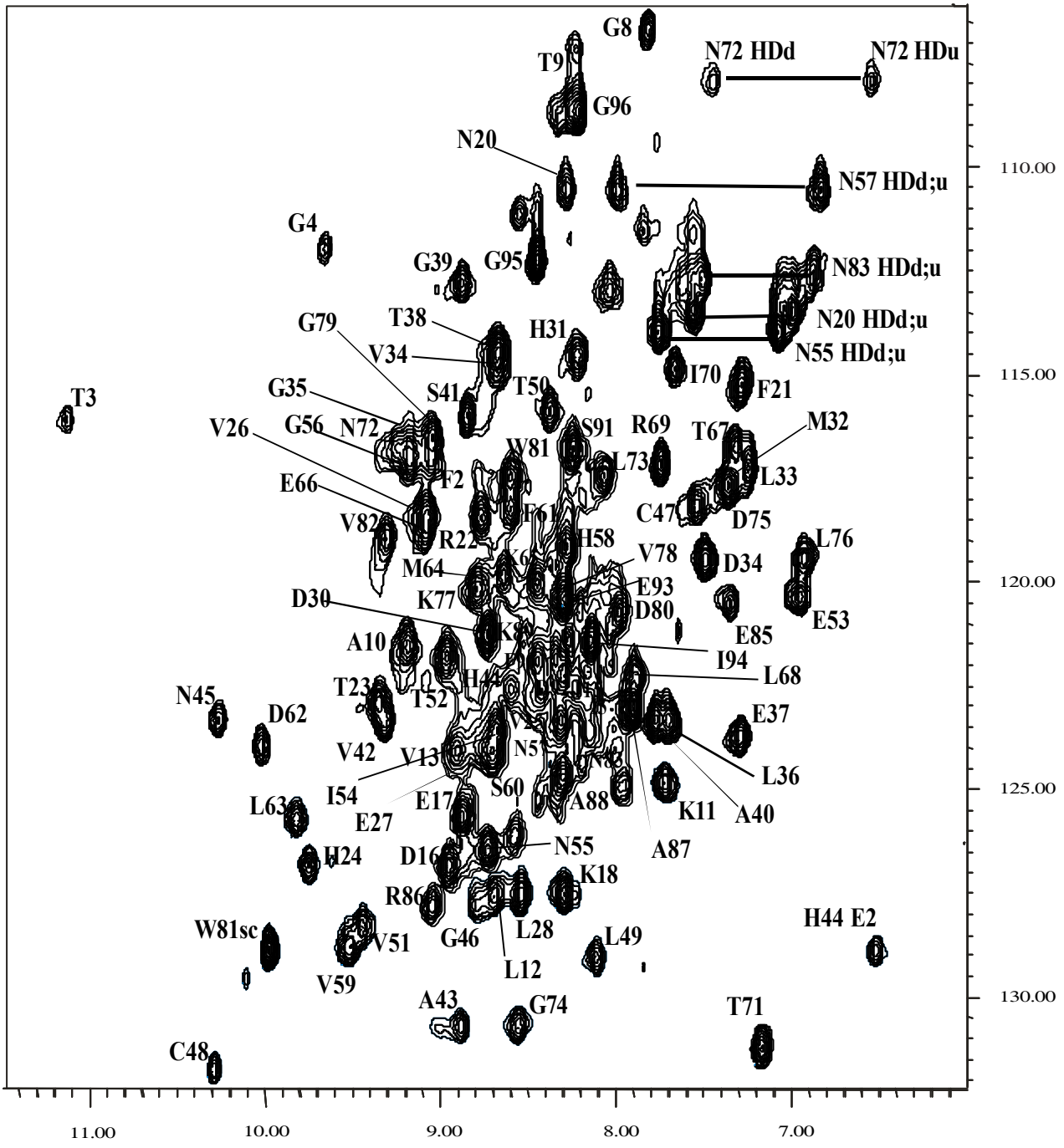


Fig. 3.8: *Fully assigned ^{15}N-HSQC of RiSy-N. HN groups of some side-chains are also shown.*

3.4 Secondary Structure

The secondary structure of the RiSy monomer was determined using C_α, C_β, C´ and H_α chemical shift indices [88, 89] (Fig. 3.9), together with backbone NOE patterns (Tab. 2.4) and $^3J_{HN}H_\alpha$ (Fig. 2.1) coupling constants.

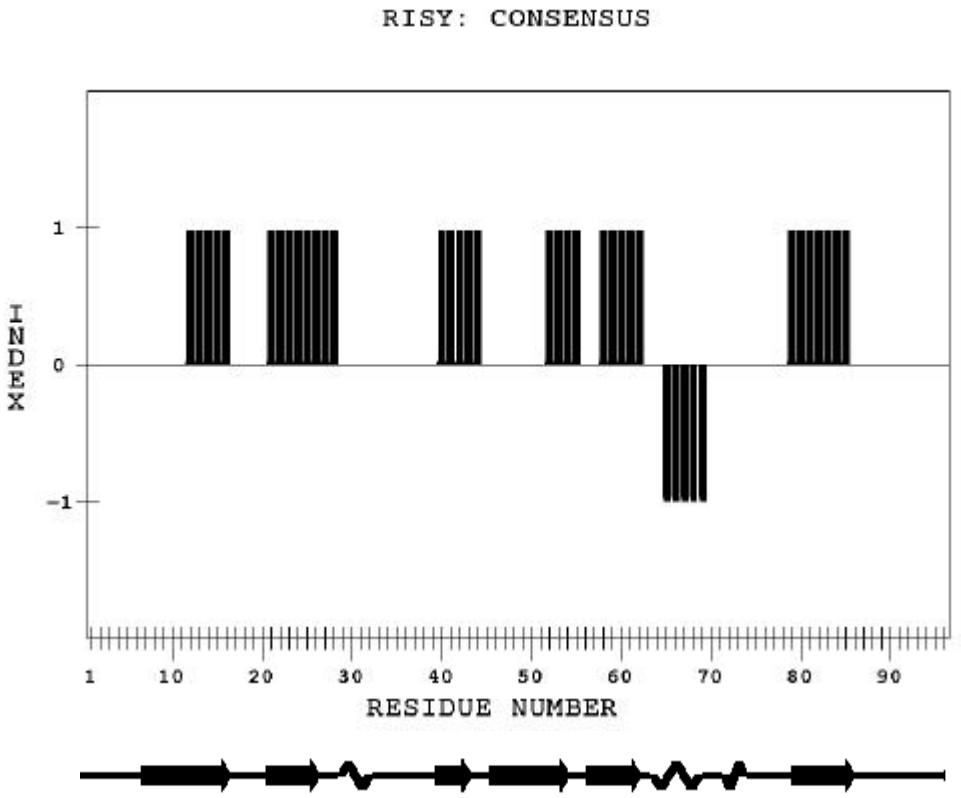
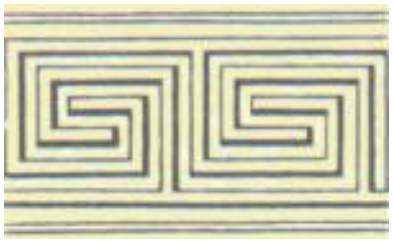


Fig. 3.9: *Secondary structure prediction with CSI. Consensus rates (+1 : β-sheet and –1: α-helix) are derived from C_α, C_β, C' and H_α chemical shifts. The first helix can't be predicted from CSI values.*

Six β-strands and one α-helix were clearly predicted from these data. The folding topology of the protein was then determined by consideration of NOE contacts between the β–strands, as shown in Fig. 3.10. Strands β2-β5 adopt a "greek key" fold, with two antiparallel flanking strands arranged to form a β-barrel.



The Greek-key or meander (shown aside) describes an ornamental motif which is used in architecture as a trim line. The endless pattern is bent orthogonal in periodic distances. It first appeared in the early stone age to ornament ceramics and was later rediscovered and extensively used by the Greeks. During classicism the pattern was very popular and underwent many modifications. In some languages, it is also known as “running dog”. The name meander is derived from the strongly curved river “maiandros” situated in Asia Minor. The greek-key fold is also well-known as a super-secondary structure. One side of an antiparallel β-sheet typically faces the solvent, and the other faces a hydrophobic core. Loops and α-helices in these structures are almost always in contact with the solvent and link opposite ends of the sheets.

While helix α2 was clearly predicted by the secondary structure data, a putative, second helix (α1) was less clearly indicated (Fig. 3.9). This helix was later confirmed during quantitative structure determination and the fold can thus be viewed as two almost identical repeating ββαβ units; for example, note that the corresponding strands of both units are of approximately equal length and that β-bulges occur at equivalent positions in the β1 and β4 strands (Fig. 3.10). Some sequence similarities can be detected between the two repeat units; however, these are both limited and weak.

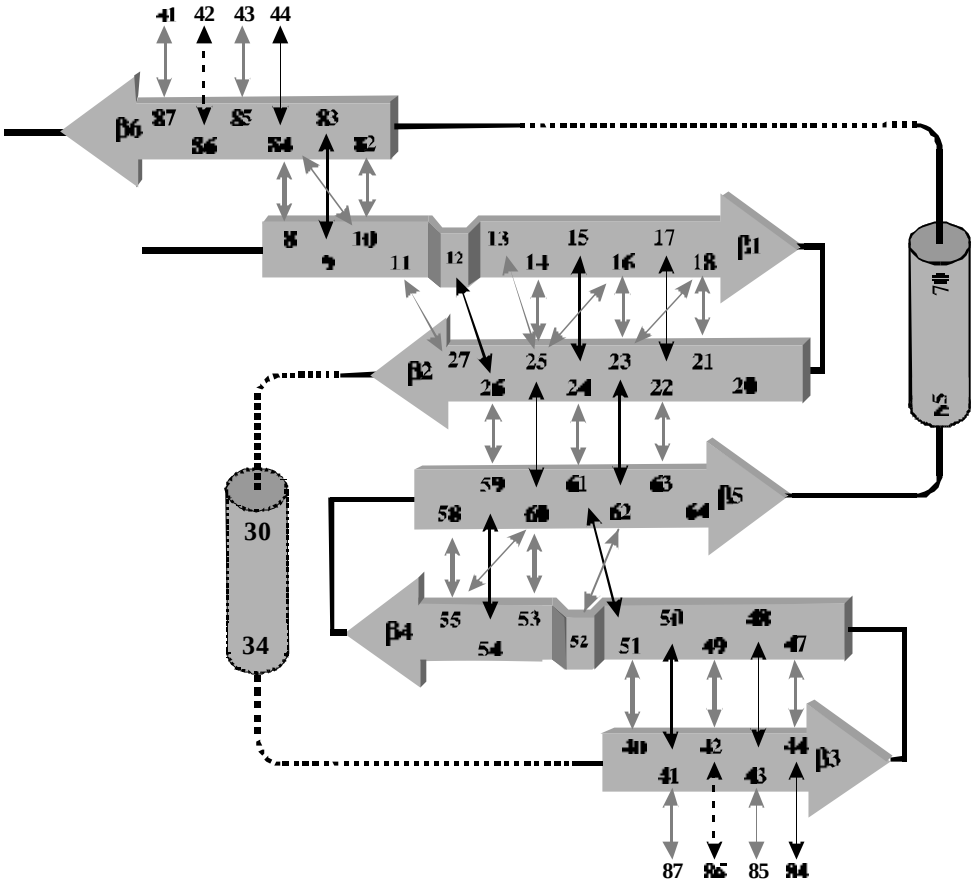


Fig. 3.10: *The topology of RiSy-N. NOE contacts defining the β-strands are indicated (H^N-H^N, light arrows; H_α-H_α, dark arrows). RiSy-N forms a six-stranded "greek-key" β-barrel. Helix α1 (broken) was only weakly predicted by secondary structure indicators and is somewhat flexible.*

3.5 Tertiary Structure

3.5.1 Structural data

The assignment of NOESY connectivities was completed using a strategy involving the combination of complementary 3D NOESY experiments. The strategy [102] involves the acquisition of four hetero-edited 3D NOESY spectra based on the HSQC-NOESY-HSQC experiment in addition to standard 3D ^{15}N-HSQC-NOESY and 3D ^{13}C-HSQC-NOESY experiments. We name these spectra the CNH-, NNH-, CCH- and NCH-NOESY, according to the three shifts recorded (see chapter 2.2.5). These experiments exploit the higher dispersion of the heteronuclei over protons to minimise ambiguities. In the case of the "orthogonally filtered" CNH- and NCH-NOESY experiments [102] diagonal signals and water exchange cross-peaks, which could contribute to overlap, are completely absent.

An HNH-NOESY and an NNH-NOESY both using water flip-back techniques were recorded on the ^{15}N-labelled sample, and an HCH-NOESY, a CCH-NOESY, an NCH-NOESY and a CNH-NOESY on the double-labelled sample. The mixing times of all these experiments were set to 80 ms. NOE cross-peaks in the 3D spectra were converted into distance ranges after rescaling according to corresponding HSQC intensities. Cross-peaks were divided into four classes, strong, medium, weak and very weak, which resulted in restraints on upper distances of 2.7, 3.2, 4.0 and 5.0 Å, respectively. Lower distance bounds were also applied for restraints where the maximum possible separation was below 5 Å, using minimum distances of 2.7 Å and 3.2 Å, for weak and very weak peaks, respectively. Allowances for the use of pseudo- atoms (using r^{-6} averaging) were added where necessary.

Torsion angles restraints were included for backbone ϕ-angles within elements of secondary structure based on $^{3}J_{HN_{H\alpha}}$ coupling-constants ($J > 8$ Hz: -120 ± 30^{o}, $J > 9$ Hz: -120 ± 20^{o}, $J < 6$ Hz: -60 ± 30^{o}, $J < 5$ Hz: -60 ± 20^{o}).

Dihedral restraints were also applied for the side-chain χ_1 angles of the 49 residues for which predominant rotamers were identified during stereospecific assignment; these included all threonine and valine residues, four of five isoleucines and both prolines. The tolerance used was $\pm 45^{o}$, with the exception of proline residues where the predominant ring pucker was determined from the χ_1 angle and a restraint of 20 ± 20^{o} or -20 ± 20^{o} applied as appropriate.

The combination of these spectra provided a large number of unambiguous NOEs which allowed the calculation of an initial model structure for the RiSy monomer. This model structure was then used in an iterative procedure by which model structures of increasing resolution were used in the assignment of further NOEs.

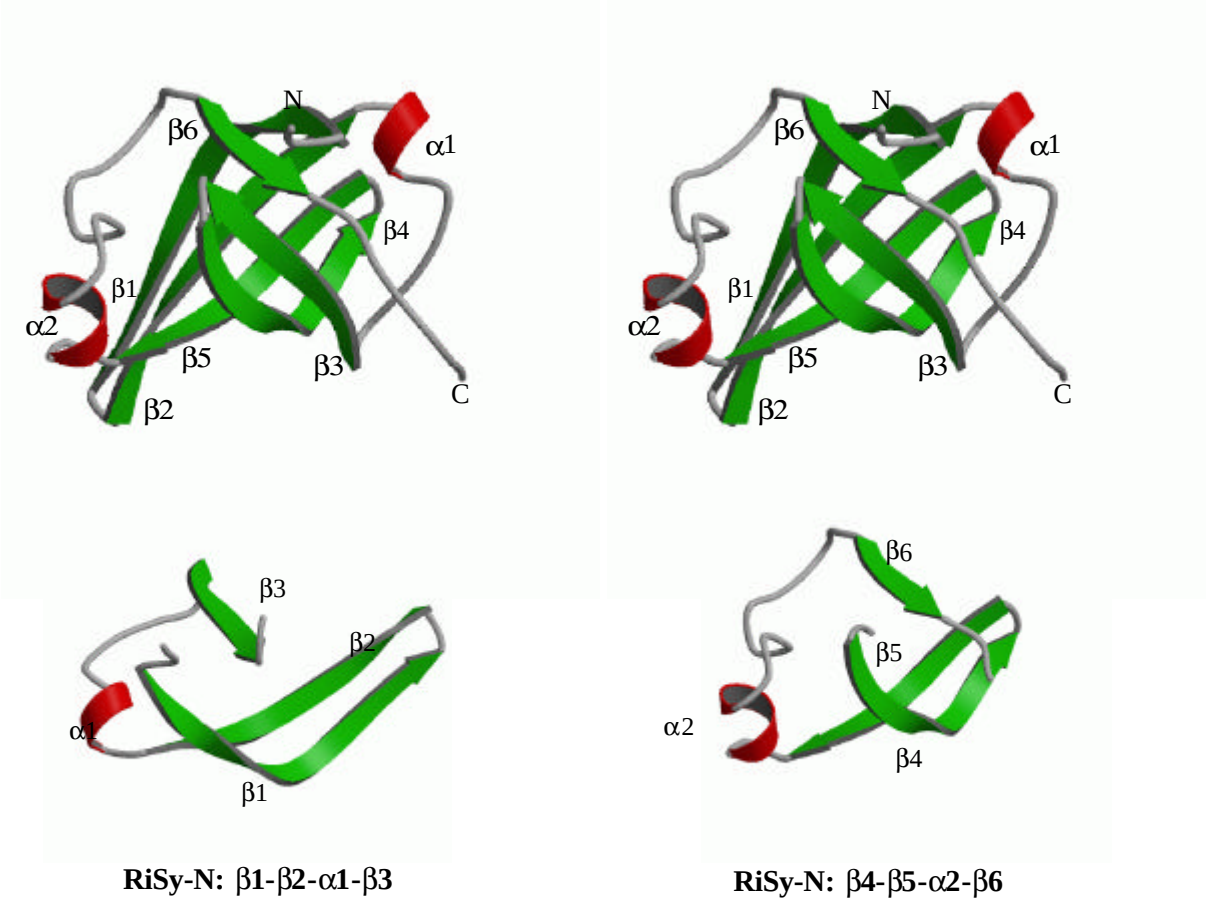
3.5.2 Structure and structure calculation

The monomer structure was calculated in XPLOR using standard simulated annealing protocols [162] which yielded sets of 50 structures. These were refined via a second simulated annealing step using a non-bonded function supplemented with a conformational database potential [171, 172] which directs structures toward conformations which are within empirical limitations generated from protein structure databases. Subsets of structures were selected on the basis of low deviations from experimental restraints and an average structure calculated and regularised in order to obtain a representative structure.

The upper panel of Fig. 3.11 shows the RiSy-N monomer to be a typical six-stranded β-barrel. Two antiparallel β-sheets with almost orthogonal chain directions are linked by the longer β1 and β4 strands. The β-bulges in these strands provide the twist necessary to close the barrel, while the two α-helices cover its open ends.

The overall fold shows a degree of symmetry reflecting the symmetry in its secondary structure (Fig. 3.10). For example, β-strands β1-β3 can be superimposed over β4-β6 with a backbone RMSD of 2.5 Å (Fig. 3.11, lower panel). Clear differences can be seen, however, in the two helices. Helix α1 was only weakly predicted in the secondary structure data, and high amide proton

water exchange rates show it to be somewhat unstable (Fig. 3.12). In contrast, helix α2 is more stable, as indicated by the low water exchange rates, and shows numerous NOE contacts to strands β2 and β4 . Another small helix or helical turn immediately follows α2 in the β5-β6 loop, a feature notable in the comparison of the RiSy-N structure with that of structurally related proteins.



RiSy-N: β1-β2-α1-β3 **RiSy-N: β4-β5-α2-β6**

Fig. 3.11: *Upper panel: stereoview of the RiSy-N monomer. The structure displays a typical greek-key scheme where both ends of the barrel are closed by an α-helix. Lower panel: RiSy-N sub-domains reflect the high symmetry of the protein with respect to the two halves of the monomer.*

3.5.3 Structure comparison

The lack of significant sequence similarity to proteins of known structure has previously precluded homology-based structure predictions for the riboflavin synthases. However, structure-based searches using the RiSy-N monomer

calculated here reveal some interesting similarities which support the mode of riboflavin binding, as discussed in chapter 3.6. The most similar structures found in searches using DALI [173] are within the flavin-adenine dinucleotide binding domains of the ferrodoxin reductase family, e.g., flavodoxin reductase [174] (PDB code 1FDR) (Fig. 3.12), NADPH-cytochrome p450 reductase [175] (1AMO), and flavohemoprotein [176] (1CQX) or the flavin mononucleotide binding site of phthalate dioxygenase reductase [177] (2PIA). Significantly, all of these proteins bind the flavin in a binding pocket almost identical to that observed for RiSy-N, and with the flavin rings and ribityl chains in similar orientations. The consensus binding mode of the group involves the hydrogen bonding of the flavin ring to backbone acceptors and donors of two adjacent β-strands (Fig. 3.12b).

The most striking structural similarities between RiSy-N and the related flavoproteins involve their β-barrels which share identical greek-key folds. In contrast, the helices are rather variable. Most members of the group lack the RiSy-N α1 helix, their β2-β3 loops being unstructured. The α2 helix also differs in length, position and orientation, with RiSy-N being one of the most divergent examples. Whereas the RiSy-N helix begins immediately after the end of β5 (at K65) with its axis roughly perpendicular to the strand direction, the helix in the related proteins begins later with its axis roughly antiparallel to β5 as outlined for the particular case of FAD (Fig. 3.12b). As mentioned earlier, RiSy-N contains a short helix or helical turn over residues L73-D75. This turn occupies exactly the same position at the end of the helices in the related proteins. Seen in this light, α2 and the succeeding helical turn in RiSy-N seem to form a sharply kinked version of the related helix. It is possible that inter-domain contacts could be responsible for this difference. For example, note that the kink in the helix is located at the inter-domain interface and close to the side-chain of N83 of the dimeric partner (Fig. 3.20). It is also possible that the difference is induced by ligand binding, although this is not supported by chemical shift differences between liganded and

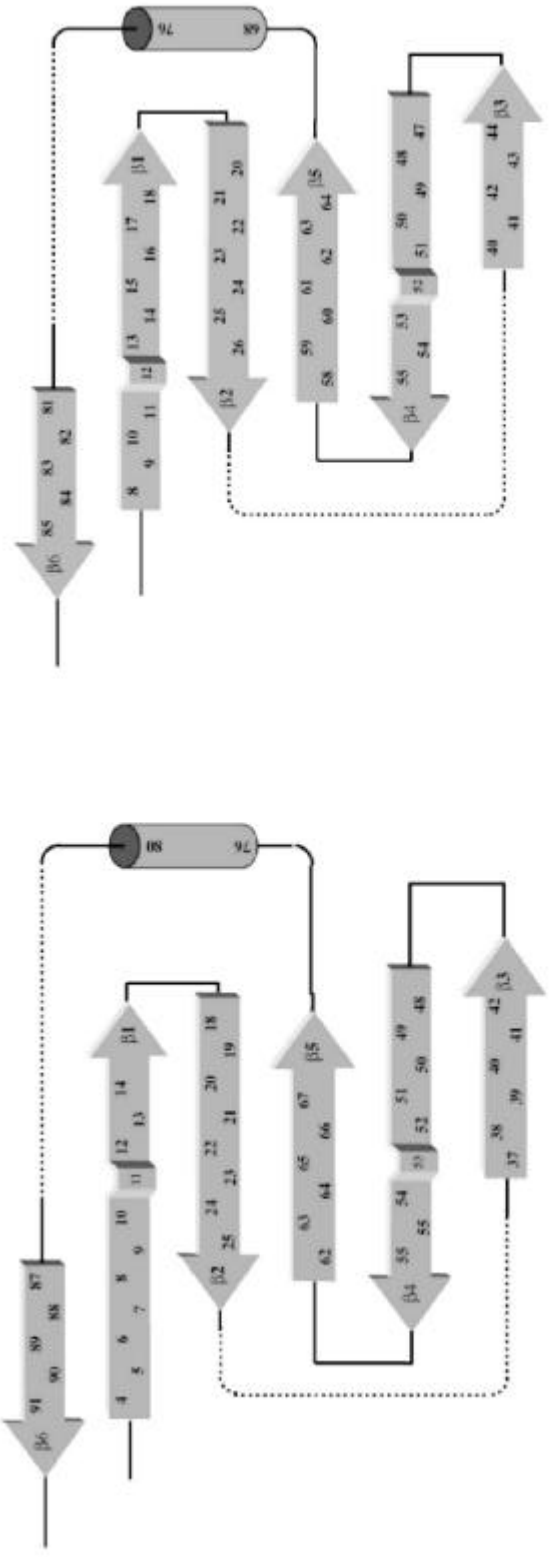


Fig. 3.12.a: *Both, the topologies of RiSy-N (right side) and of the N-terminal domain of Flavodoxin Reductase (left side), form a six-stranded "greek-key" β-barrel. The helix α1 is omitted in the RiSy-N topology to point out the similarities between both proteins.*

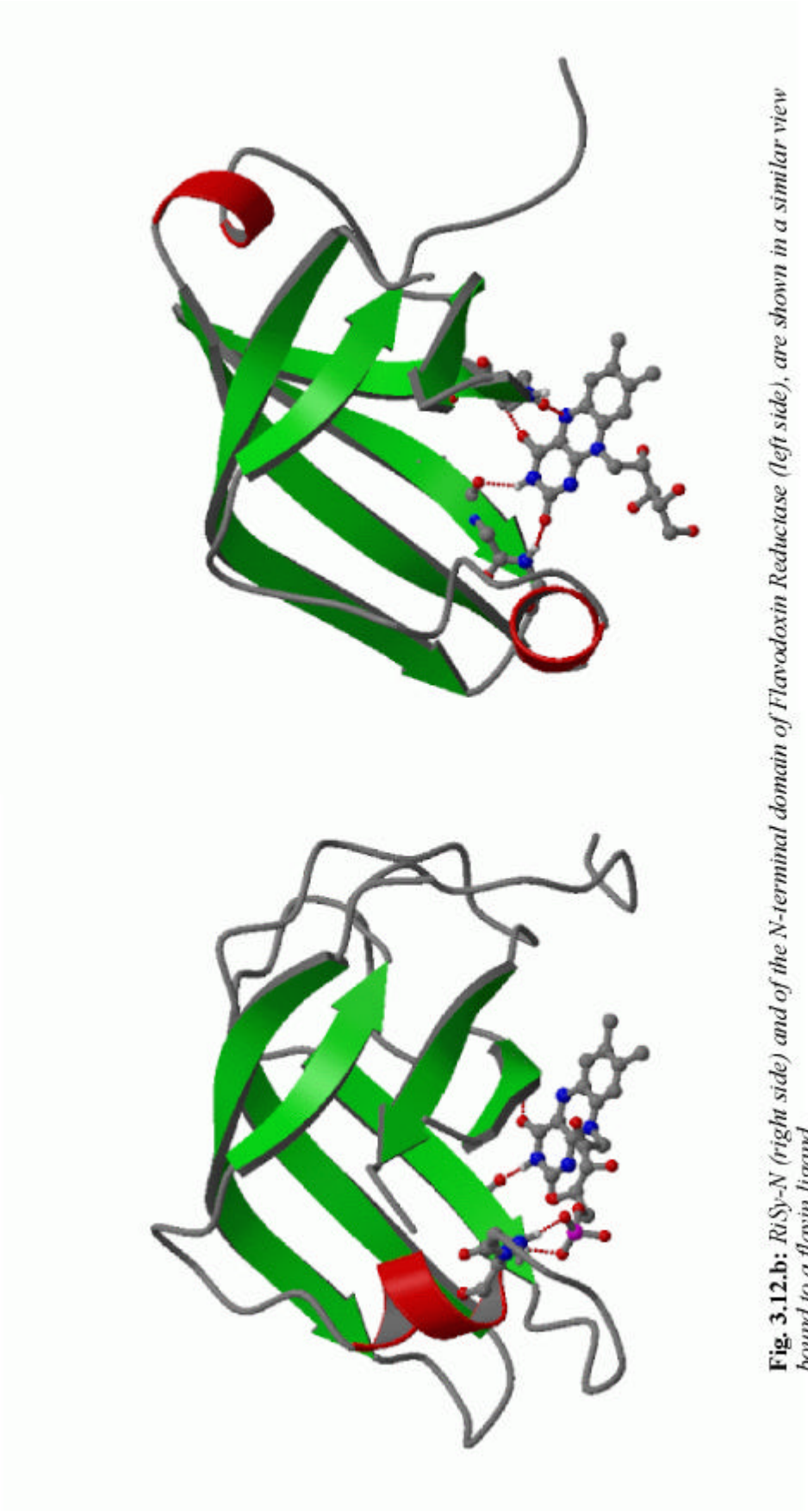


Fig. 3.12.b: *RiSy-N (right side) and of the N-terminal domain of Flavodoxin Reductase (left side), are shown in a similar view bound to a flavin ligand.*

unliganded forms which are only minor in this region. However, in the present example,we can remark that its bound flavin molecule is rather different to that in RiSy-N. Especially the phosphate group involves two hydrogen bonds with flavodoxin reductase and thus contributes in a large part to the definition of the binding mode of the ligand. It can then be viewed that the ligand influences the position of the α2-helix along the flexible loop.

3.6 Quaternary Structure

3.6.1 Structural data

Cross-labelled sample and mass spectroscopy

During the assignment of the NOESY spectra, several contacts could not be rationalised with the structure of the monomer as mentioned in chapter 3.5. These contacts were also considered as potential inter-molecular contacts. Several attempts were made to confirm these contacts using a sample containing mixed singly ^{15}N- and singly ^{13}C-labelled components, where heat denaturation was used in an attempt to dissociate the homodimers (cf. chapter 2.2.6: isotope filter experiment). A 0.5 mM cross-labelled sample was expected from the statistic recombination of the mixture.

The nature of the resulting sample was controlled using mass spectroscopy. The experiments were run in the group of Prof. Dr. J. Buchner, by Helmut Krause. Three MALDI-TOF experiments were acquired on a Bruker Biflex-3 spectrometer in linear modus, using an α-cyanohydroxy-cinnamic acid matrix. The first two spectra permit to control the molecular mass of the [U-^{15}N]-labelled dimer and of its [U-^{15}N, ^{13}C] homologue. Also, the spectra were calibrated on a test sample using lysozyme that possesses a well-known molecular mass of 14 kDa.

A mass of 21408 and 22301 amu were obtained for the singly and double labelled test samples, respectively, which is in perfect agreement with the calculated masses of 21350 and 22272 amu (Fig. 3.13a, Fig. 3.13b). The spectra of the mixed sample shows a quintett (Fig. 3.13c). The three larger masses were easily assigned to the ^{13}C-dimer (22055 amu), the mixed dimer (21724 amu) and the ^{15}N-dimer (21400 amu). It appears that some fragmentation occurs due to the presence of two additional peaks for which the molecular mass was around 21 kDa. However, the fragmentation is independent of the isotopic labelling and the relative intensity of the three main signals can be compared. It results that the mixing occurs (all signal show a similar intensity).

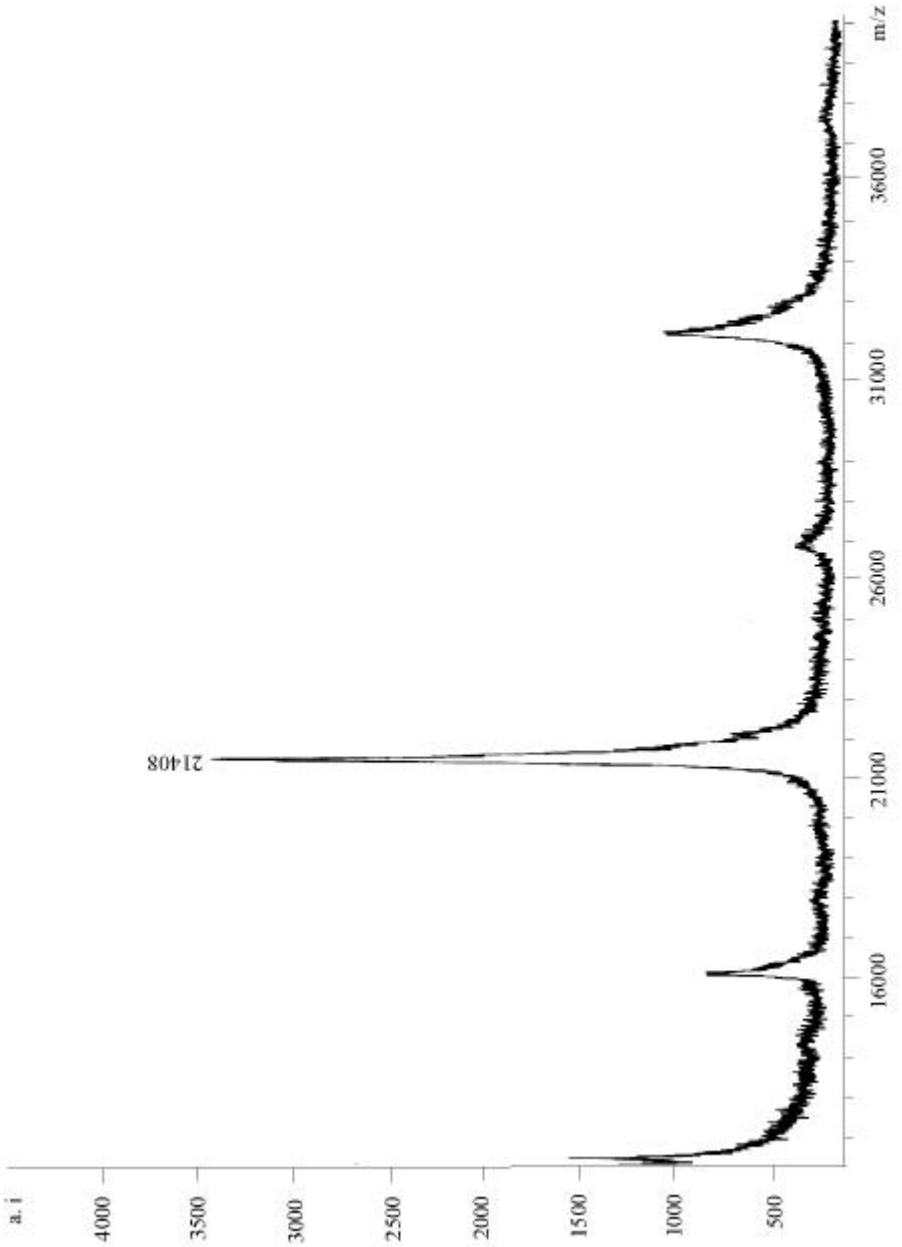


Fig. 3.13.a: *MALDI-TOF spectrum of the [U-^{15}N] labeled RiSy-N dimer. A mass of 21350 amu is expected.*

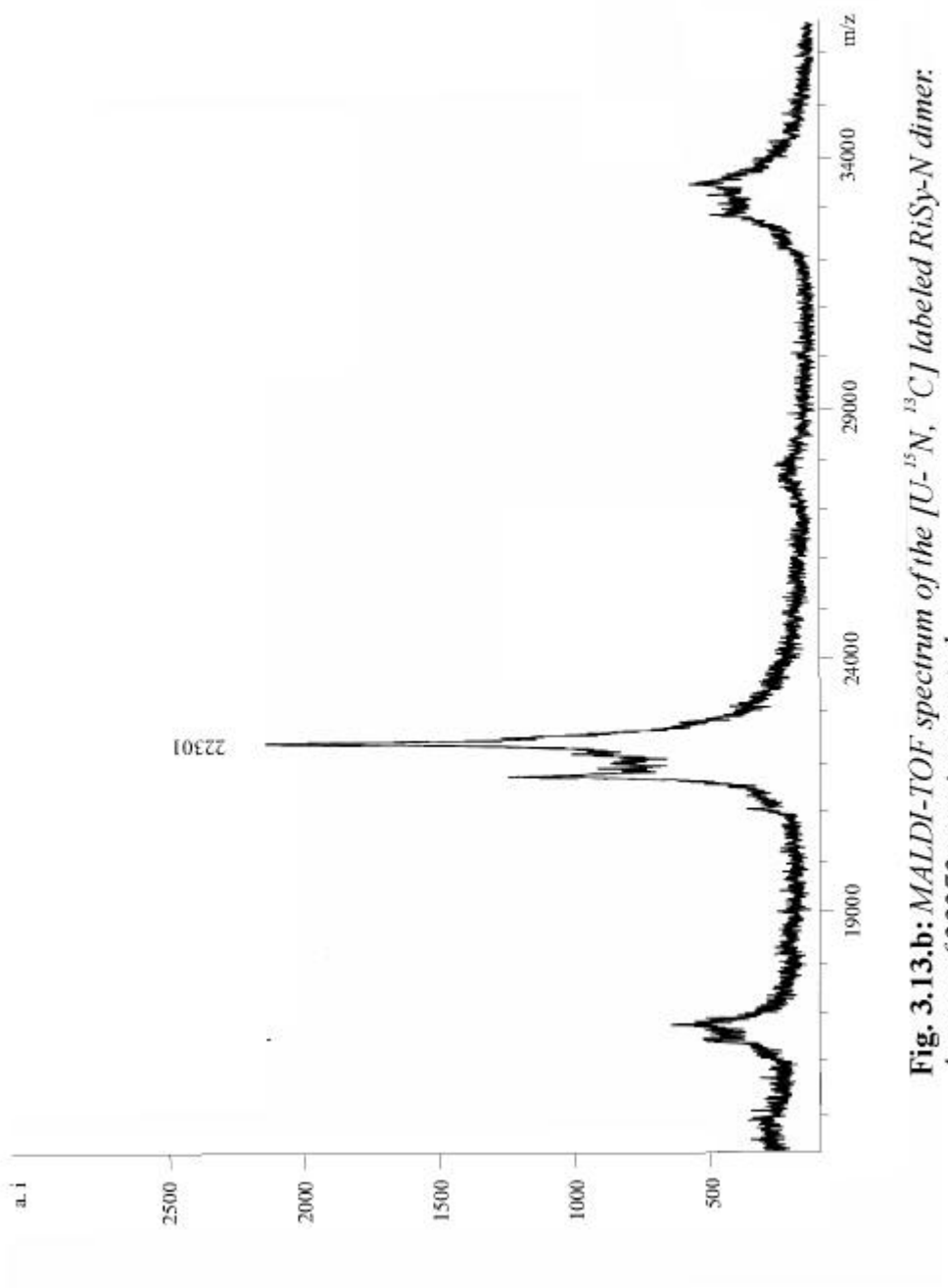


Fig. 3.13.b: *MALDI-TOF spectrum of the [U-^{15}N, ^{13}C] labeled RiSy-N dimer. A mass of 22272 amu is expected.*

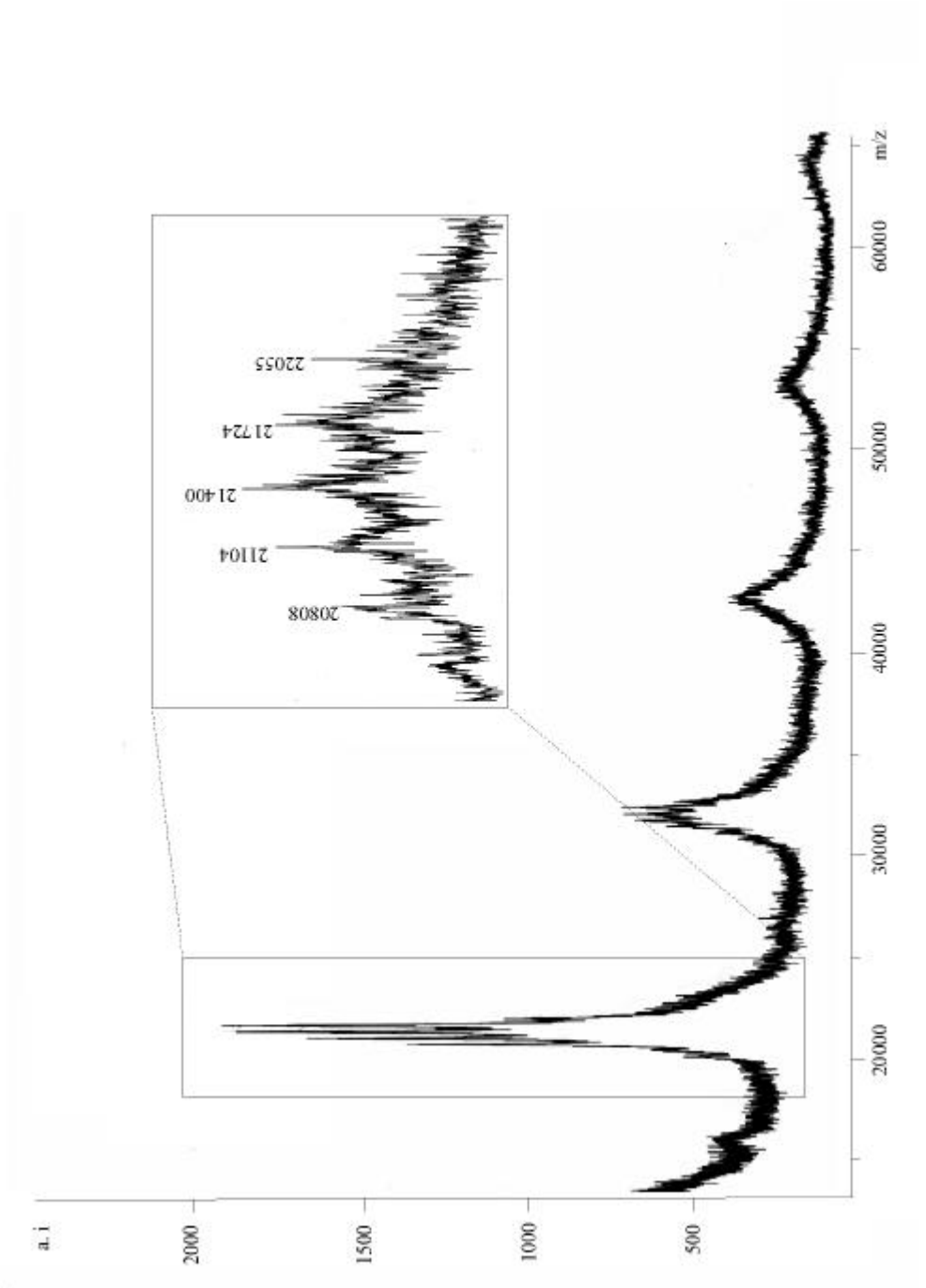


Fig. 3.13.c: *MALDI-TOF spectrum of the cross-labeled RiSy-N sample: a triplet was expected from the statistic recombination of the mixture. The calculated masses are 22014, 21682 and 21350 amu for the [U-^{15}C] dimer, the mixed sample and the [U-^{15}N] dimer, respectively.*

However, the concentration of the cross-labelled dimer is expected to be twice the one of the singly labelled dimers if the sample is produced with an initially identical concentration and if the mixing occurs completely. The intermediate intensity observed for the orthogonally labelled sample and the very small one for the singly ^{13}C-sample reflect the concentration dependence of the dimer on the starting proteins, and thus the difficulties in producing a mixed sample, even under denaturing conditions.

As a result, no correlation were observed in either CNH-NOESY or various filtered-edited NOESY spectra run on this sample, even in a 2D version of the former run with 1024 scans per slice. Thus, it appeared that mixing of the components takes place in a minimal proportion.

Water exchange experiment

MEXICO experiments [123, 178] using six mixing times (50, 100, 150, 200, 250 and 300 ms) were used to determine rates of exchange of H^N protons with water. The results are reported Fig. 3.14. Protons with low exchange rates were considered to be used as H-bond restraints where H-bond acceptors were consistently suggested in preliminary calculations. In this study a new approach was employed where H-bonds were treated as normal covalent bonds added via „patches“ to the XPLOR molecular structure. Parameters were included for the length and angle of the bond (2.12 Å with force constant 14.0 kcal/mol/Å^2 and 0° with force constant 4.0 kcal/mol/rad^2, respectively). We have found this approach to give favourable H-bond geometries when compared to conventional methods employing two distance restraints per H-bond, while maintaining a distribution of geometries comparable to that seen in related crystal structures. To avoid over-interpreting the exchange data, no H-bond restraints were applied outside the elements of secondary structure.

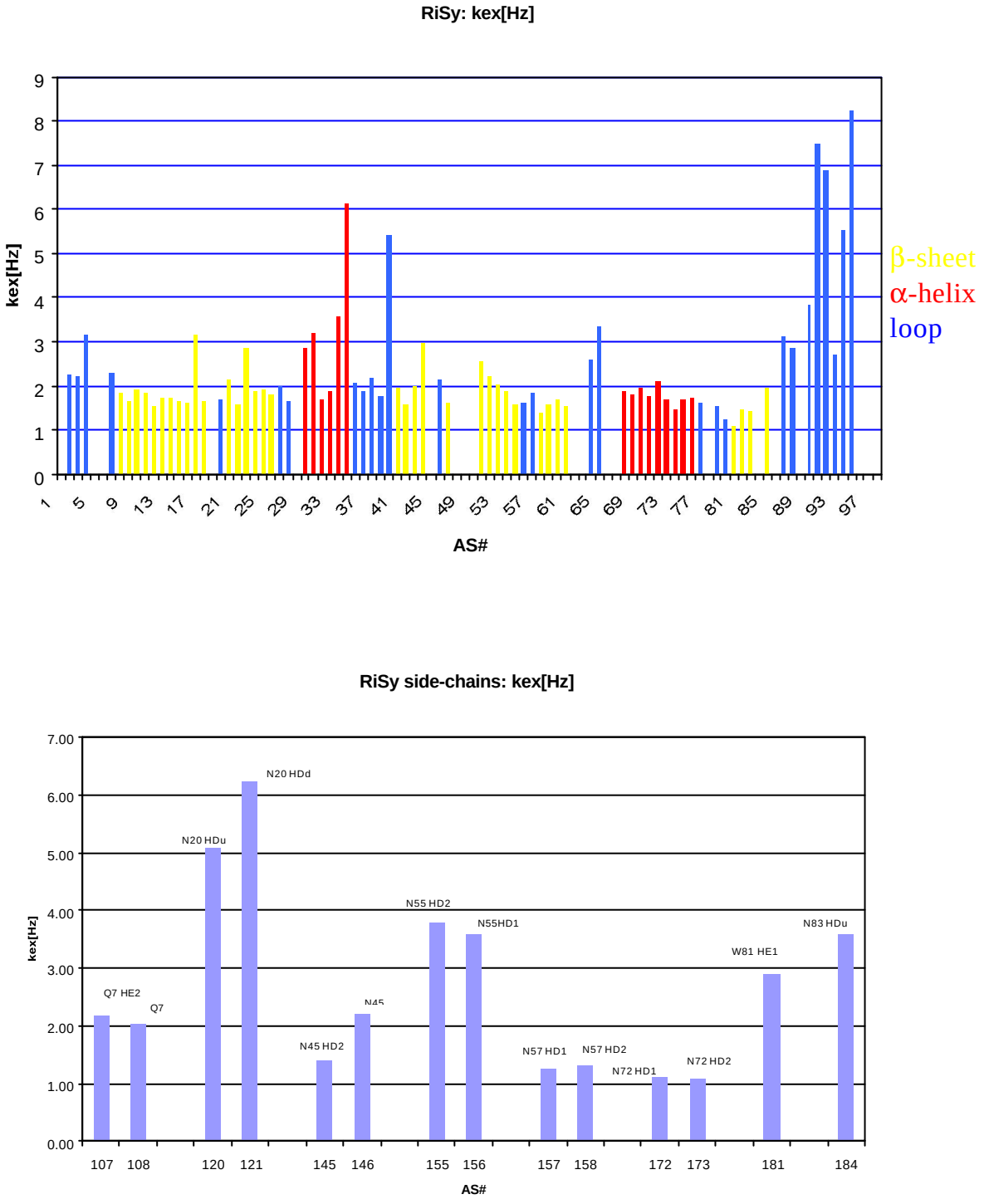


Fig. 3.14: *Backbone (upper panel) and side-chains (lower panel) water exchange rates are derived from MEXICO experiments measured with six different mixing times. Low (< 2.2 Hz) exchange rates for the N-termini (M1-V6) and side-chains amide groups 72 and 45 are consistent with their involvement in hydrogen bonding in the definition of the binding site and/or the site of dimerisation.*

Heteronuclear NOE

The amplitude of motions on the pico- to nanosecond time-scale were measured on a [U-^{15}N]-labelled RiSy-N sample with a standard ^{15}N{^{1}H} heteronuclear NOE experiment. The experiment was run at 750 MHz with a 3 s proton presaturation time (appendix B4). In accordance with the results of water exchange experiment, het-NOE rates of ~ 0.88 for M1-V6 confirm the implication of the N-termini in the binding and/or dimerisation site. Indeed, comparable flexibility to the one of the C-termini would be expected for the N-termini that do not present any secondary structure up to G8, and thus should have an undefined structure.

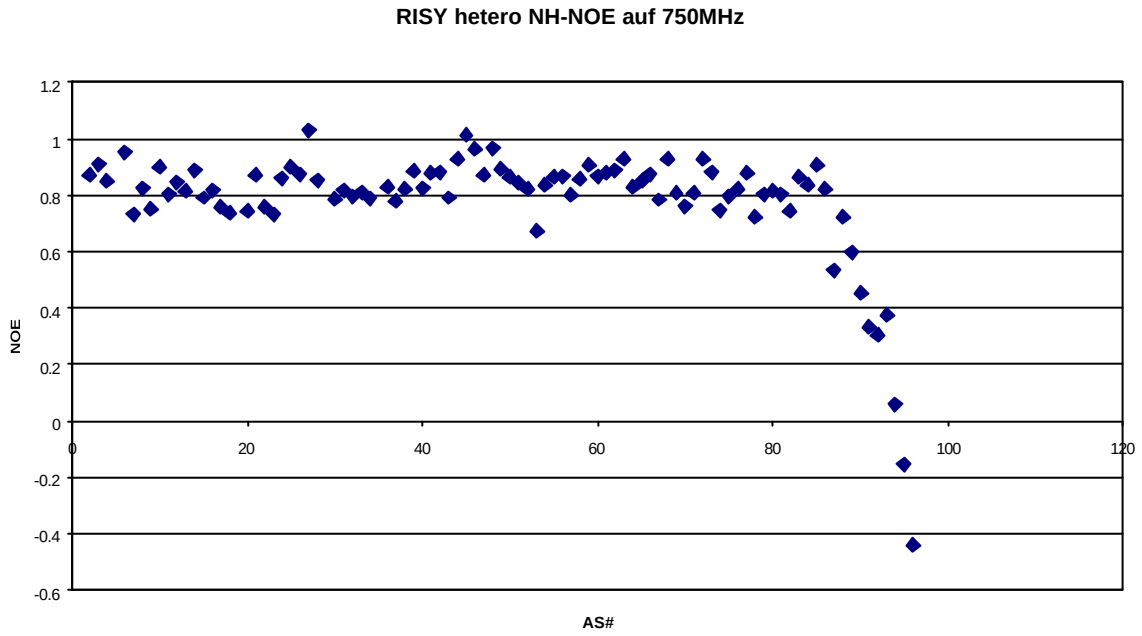


Fig. 3.15: *The heteronuclear-NOE values show two extremely different behaviour between the C- and N-termin of RiSy-N, thus reflecting the role played by the N-terminal octapeptide in the definition of the quaternary structure.*

Ligand titration

A sample containing the apo-form of the N-terminal domain of RiSy was produced in order to measure a series of riboflavin titration experiments. First a ^{15}N-HSQC was acquired on the ligand-free sample. Then, four HSQC were

recorded after addition of riboflavin in the following proportion: 10 %- 33 %- 66 %-100 % where 100 % riboflavin added means that both, the ligand and the protein, are of the same concentration. Thus, the ^{15}N-HSQC correlations of the apo-RiSy-N could be traced and their assignment was easily achieved together with the help of an HNHA spectrum that was previously measured on the ligand-free sample (appendix B5). The superposition of the apo- and holo-spectra is displayed in Fig. 3.16.

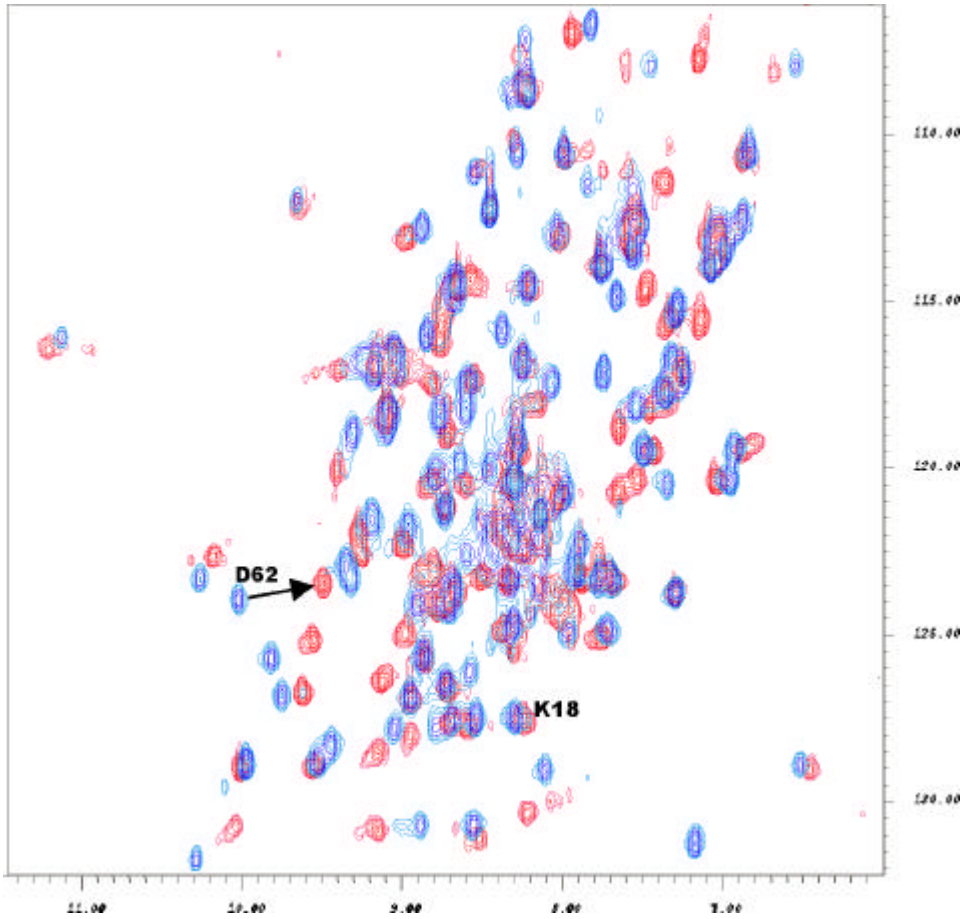


Fig. 3.16: *^{15}N-HSQC experiments measured in presence (blue spectra) or absence (red spectra) of riboflavin under similar experimental conditions. The chemical shift changes of D62 for example reflect structural changes of the protein due to the ligand. In contrast K18 and other residues are not influenced thus suggesting that the binding takes place at an extremity of RiSy-N and not in its core.*

Isotope filter experiment

The lack of unambiguous monomer-monomer contacts was somewhat compensated for by the observation of contacts between [U-^{13}C, ^{15}N]-labelled riboflavin and ^{15}N-labelled RiSy-N, as well as between unlabelled riboflavin and doubly labelled RiSy-N. The assignment of bound riboflavin was made from a ^{13}C-HSQC spectrum acquired on a sample containing ^{13}C/^{15}N-labelled riboflavin and ^{15}N-labelled RiSy-N. An isotope-filtered 2D-NOESY spectrum (Fig. 3.17, Fig. 3.22), was particularly useful in the latter case. The pulse program shown in the Fig. 3.17 was especially developed in our laboratory for this purpose.

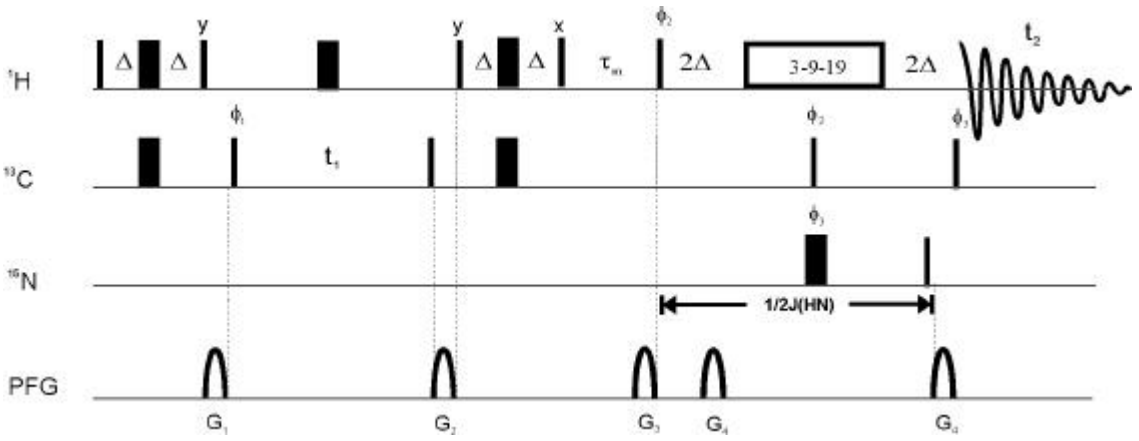


Fig. 3.17: *Pulse sequence of the ^{12}C/^{14}N-filtered 2D-^{13}C(H)-NOESY. The phase cycle is: f_1 = 0 2; f_2 = 0 0 2 2; f_3 = 0 2; f_{rec} = 0 2 2 0. D = 1/(2$\times{}^1J_{CH}$) and the 3-9-19 module represents a traditional Watergate [179] sequence.*

The sequence begins with a traditional HSQC step, thus recording the ^{13}C chemical shifts in the indirect dimension. The magnetization is then transferred back on z where the NOE mixing time of 100 ms takes place. The last part of the sequence shows an isotope-filtered module after which only the ^{12}C/^{14}N-bound protons are detected in the t_2 dimension. This latter achieves a double filtration of ^{13}C magnetization together with a single filtration of ^{15}N magnetization. Both are based on the same principles. The heteronuclear magnetization is send into unobservable double quantum coherence, either via an HMQC-step in case of undesired carbon magnetization or via an INEPT polarization transfer type step for

the nitrogen magnetization. Additionally, in order to prevent intensity loss due to extensive transversal relaxation, the filter sequence is optimised by concatenation with the watergate module for solvent suppression. It allows the possibility to implement two successive HMQC-type filter-schemes with delays set to two different values for a better suppression of the ^{13}C-bound protons (see chapter 2.2.6: isotope filter).

Several filtering orders are conjugated for improvement of the spectra quality. Indeed, residual diagonal signals that remain in the transversal plane will be defocused by the second pulsed field gradient (traditionally used to refocus the desired magnetization after suppression of the solvent) and thus won't be detected in t_2. As a final point, a destructive phase cycle (Fig. 3.17) substracts the undesired coherence while acquiring the desired one in an additive way. This experiment also combines most of the filtering methods presenting in chapter 2.2.6.

It can be observed, unfortunately, that a lot of residual "diagonal" signals are still present in the spectra (Fig. 3.18). Also, in order to facilitate the interpretation of the spectra, the heteronuclear decoupler in t_2 was omitted. Thus, the residual "diagonal" signals appear as doublet and the cross-correlations as singlet. However, it was noted that the residual diagonal signals are principally corresponding to intense CH_2 or CH_3 groups. Their residual intensity is comparable to the one of cross-peaks thus relating the high efficiency of this isotope filtered pulse sequence.

Several (11) ligand-protein contacts were identified in these experiments which could not be rationalised within one monomer, for example contacts between both methyl groups of the xylene ring of riboflavin and the methyl groups of I5 (Fig. 3.16). The participation of both RiSy-N monomers in the binding site was thus required. These intermolecular contacts bridging both monomers were used in the calculation of model holo-structures which helped to define the dimerisation of the protein. The ensuing dimer model was then used to distinguish between inter- and intra-monomer contacts using an approach based on ambiguous distance restraints [138] where each restraint was allowed to have either an intra- or

inter-molecular assignment. Any contacts which produced persistent violations in preliminary calculations of the monomer were also investigated to exclude inter-molecular interpretations, and contacts which could possibly be both inter- or intra-molecular were omitted. Final calculations included 40 inter-monomer and 31 riboflavin-protein restraints.

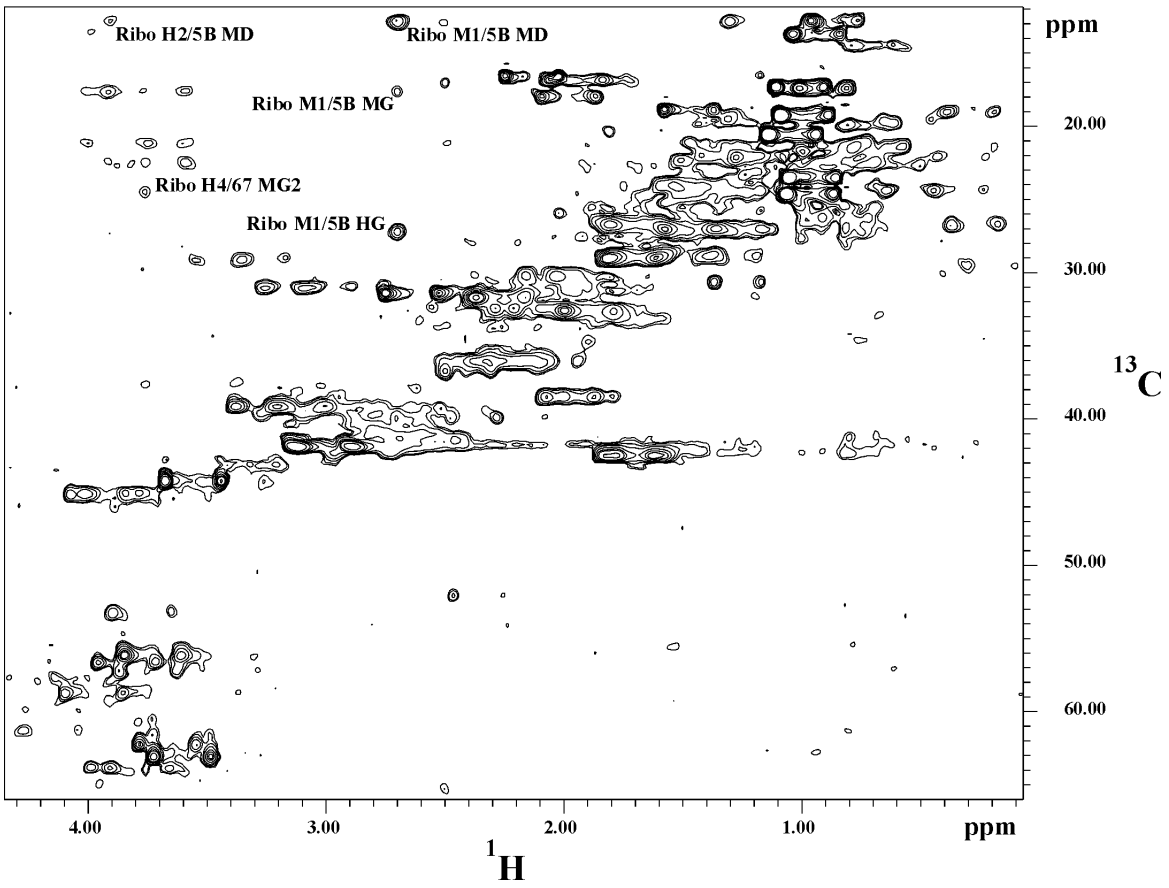


Fig. 3.18: *The $^{12}C/^{14}N$-filtered 2D-$^{13}C(H)$-NOESY spectrum. The residual diagonal signal can be easily distinguished from the cross-peaks as they appear as doublet. The intensities of the NOE cross-peaks and of the residual diagonal peaks are very close, thus reflecting the efficiency of the isotope filter pulse sequence.*

Tab. 3.1 summarises the experimental restraints used in structure calculations, which included NOE and coupling-constant data in initial runs, and a non-bonded energy function supplemented by a conformational database potential during refinement.

Tab. 3.1 also shows restraint violations and structure quality statistics for the final set of 21 structures, which was chosen such that no NOE restraint was violated by more than 0.20 Å and no dihedral restraint by more than 3°.

3.6.2 Structure and structure calculation

The dimer structure was calculated via a protocol identical to that used for the monomer, but using a starting-structure generated by duplication of the monomer. The symmetry relation used and the position of the riboflavin were determined in preliminary calculations. During refinement the XPLOR symmetry potential was applied over all well-defined residues (F2- D92) and riboflavin, ensuring that the two monomers are superimposable without enforcing a symmetry relation. In addition, a symmetry restraint was included for each intermolecular restraint, specifying that symmetry-related distances were identical.

Tab. 3.1 summarized the structure quality determined using the program PROCHECK [180]. It directly reflects the quality and quantity of the structural parameters used for the final structure calculations. The combination of the ϕ, ψ angles is especially representative of the quality of the structure. They are displayed in the Ramachandran diagram (Fig. 3.19) for the averaged minimized structure. 90.4 % of the residues (from which M1, the glycines and prolines are excluded) are in the most favoured regions, 9.6 % in the additionally allowed regions and none in either the generously allowed or the disallowed regions.

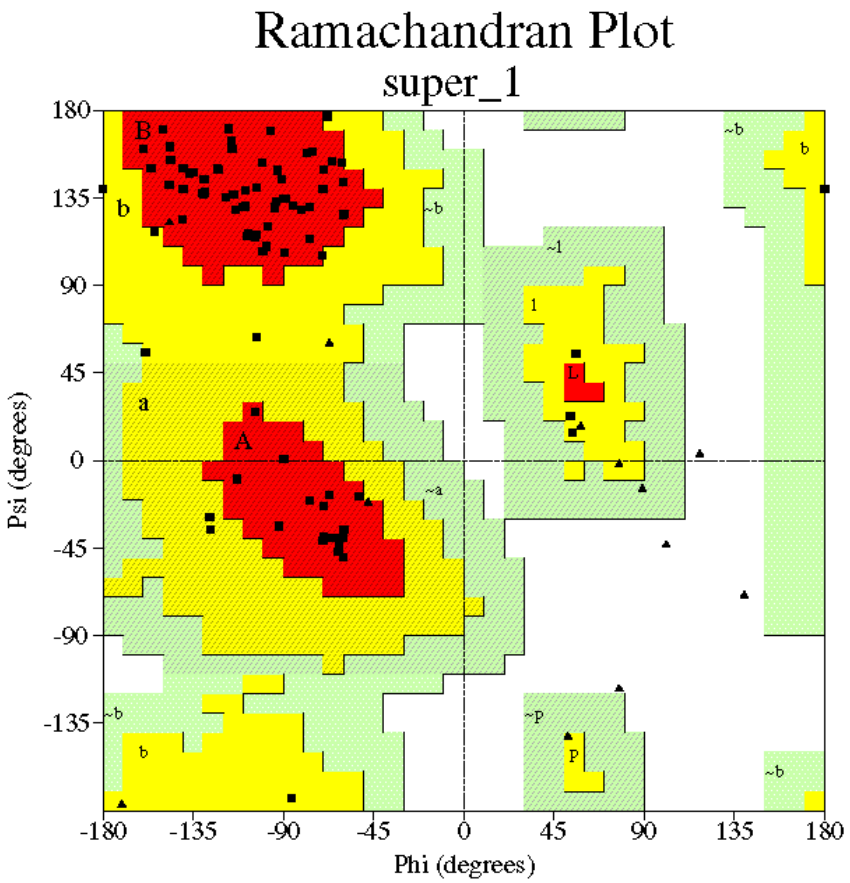


Fig. 3.19: *Ramachandran plot for the averaged minimized NMR structure of RiSy-N, determined by the program PROCHECK [180]. The regions color is: red, most favoured regions; yellow, allowed regions; green, generously allowed regions and white disallowed regions. All non glycine amino acids are represented by black squares. Glycine amino acids for which this classification is not representative are represented by black triangles. The privileged region A corresponds to helical structures, B to* ***b****-sheets and L to left-handed helices.*

Tab. 3.1: *Structural statistics and atomic R.M.S. deviations.*[a]

A. Structural statistics

R.M.S.D from distance restraints (Å)[b]	SA	$\langle SA \rangle_r$
all (2569)	0.017 ± 0.001	0.016
intra-residue (353 × 2)	0.011± 0.003	0.010
inter-residue sequential (365 × 2)	0.014 ± 0.004	0.013
medium range (159 × 2)	0.020 ± 0.009	0.020
long range (372 × 2)	0.017 ± 0.006	0.017
inter-molecular (71)[c]	0.042 ± 0.033	0.038
R.M.S.D from dihedral restraints (deg) (56 × 2)	0.419 ± 0.055	0.475
H-bond restraint violations (Å/deg)[d] (42 × 2)	1.99 ± 0.44 / 23.3 ± 11.6	2.00 ± 0.44 / 23.3 ± 10.4
Deviations from ideal covalent geometry		
Bonds (Å × 10^{-3})	2.92 ± 0.006	2.94
Angles (deg.)	0.560 ± 0.005	0.564
Impropers (deg)	1.69 ± 0.05	1.69
Structure quality indicators[e]		
Ramachandran Map regions (%)	86.1 / 13.9 / 0 / 0.0	90.4 / 9.6 / 0.0 / 0.0
Bad contacts per 100 residues	9.4	7.7

B. Atomic R.M.S. differences (Å)[f]

	SA *versus* <SA>		SA *versus* $\langle SA \rangle_r$	
	Backbone	All	Backbone	All
All residues (dimer)	1.34 ± 0.27	1.63 ± 0.28	1.70 ± 0.41	2.11 ± 0.40
Sec. structure elements (dimer)	0.27± 0.06	0.64 ± 0.06	0.36 ± 0.06	0.88 ± 0.11
Sec. structure elements (monomer)	0.19 ± 0.04	0.61 ± 0.05	0.28 ±0.05	0.84 ± 0.10
<SA> vs $\langle SA \rangle_r$[g]	0.24	0.60		

[a]Structures are labelled as follows: SA, the set of 22 final simulated annealing structures; <SA>, the mean structure calculated by averaging the coordinates of SA structures after fitting over secondary structure elements; $\langle SA \rangle_r$, the structure obtained by regularising the mean structure under experimental restraints. [a]Numbers in brackets indicate the number of restraints of each type. [c]Comprised of 31 ligand-protein and 40 inter-monomer restraints. [d]H-bonds were restrained by treating them as pseudo-covalent bonds (see structural data section). Deviations are expressed as the average distance/average deviation from linearity for restrained H-bonds. [e]Determined using the program PROCHECK [180]. Percentages are for residues in allowed/additionally allowed/generously allowed/disallowed regions of the Ramachandran map. [f]Based on heavy atoms superimpositions. Secondary structure defined as residues 5-86. [g]Based on superimposition over secondary structure elements.

The RiSy-N dimer

During the assignment of the NOESY spectra, several (18) contacts which could be assigned on the basis of unambiguous chemical shifts could not be rationalised with the structure of the monomer. For example contacts were found between the side-chain amide group of N72 and the backbone H_α and H^N of N83. These were considered as potential intermolecular contacts. Due to the difficulties in producing samples with differentially labelled monomers, it was not possible to confirm these inter-monomer contacts with specific isotope-edited NOESY experiments (see chapter 3.6.1).

The structure of the RiSy-N dimer is shown in Fig. 3.20. The dimer interface is characterised by several asparagine and glutamine residues, many of which appear to be involved in intermolecular hydrogen bonding. In particular, the low side-chain amide proton water exchange rates of N45 and N72 are consistent with their involvement in hydrogen bonding, which cannot be rationalised in the monomer structure. Other contacts involve the N-terminal strand (M1-V6), which is highly conserved in all known RiSy sequences, and the β3 strand. The upfield chemical shift of the methyl group of A43 (-0.09 ppm), which is difficult to explain in the monomer structure, is most likely induced by ring-currents in the aromatic side-chain of F2 of the dimeric partner to which it also shows several NOE contacts.

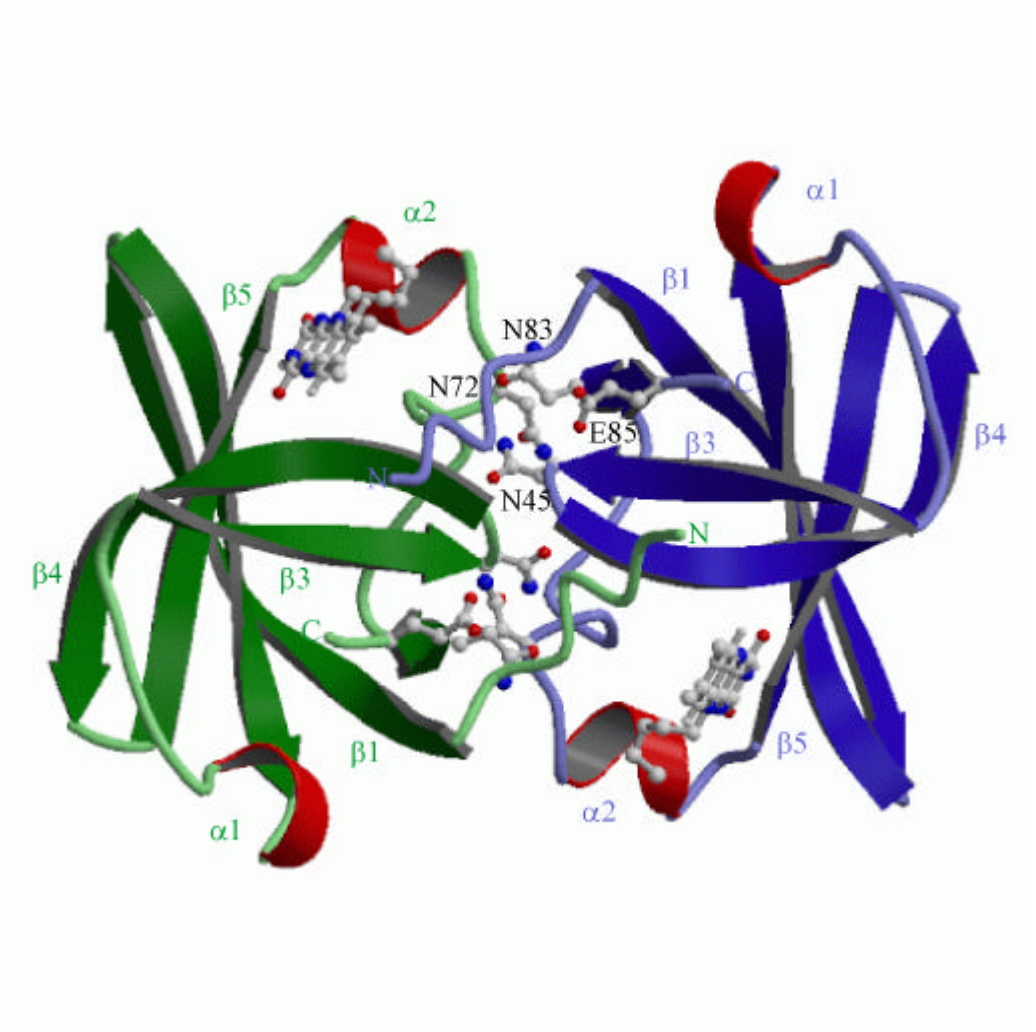


Fig. 3.20: *The RiSy-N homodimer. The view shown looks down the symmetry axis. The two riboflavin ligands are represented in ball-and-stick rendering. Also shown are the side-chains of the highly conserved residues N45, N83, E85 and N72 which are involved in inter-monomer contacts.*

Concomitantly, residues in the N-terminal strand show high positive $^{15}N\{^{1}H\}$ heteronuclear NOE values (Fig. 3.13, Fig. 3.21), rather than the low positive or negative values which would be characteristic of the free strand as in the monomer structure.

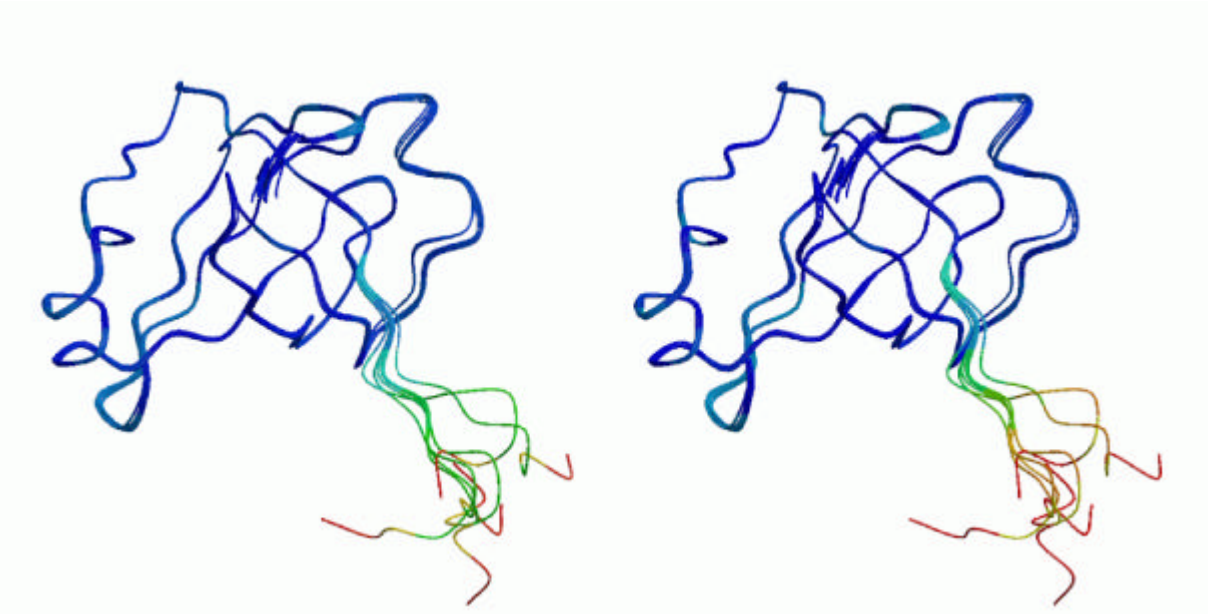


Fig. 3.21: *Stereo view of RiSy-N monomer. The final set of 21 structures is shown, superimposed over backbone atoms of secondary structure elements, as defined in Tab. 3.1. The structure is colored according to $^{15}N\{^{1}H\}$ heteronuclear NOE values (measured at 750 MHz) reflecting the amplitude of motions on the pico- to nanosecond time-scale. A linear ramp was applied from 0.2 (most flexible; red) to 0.85 (least flexible; blue). Values for four residues, I5, G35, E53 and V78, which were missing due to overlap, as well as prolines P19 and P29 were replaced by interpolation from neighbouring values. In contrast to the unstructured, flexible C-terminus, the N-terminus is relatively rigid. This figure, Fig. 3.20 and Fig. 3.22 were generated using MOLSCRIPT [181] and Raster3D [182].*

Riboflavin binding

The riboflavin binding site can be clearly located from intermolecular NOE contacts and chemical shift differences between apo- and holo-forms. The exact binding mode was calculated using 31 NOE contacts observed in various filtered/edited NOESY spectra (see above) and shows the pteridine ring to be in contact with the backbone of strands β4 and β5 (Fig. 3.22).

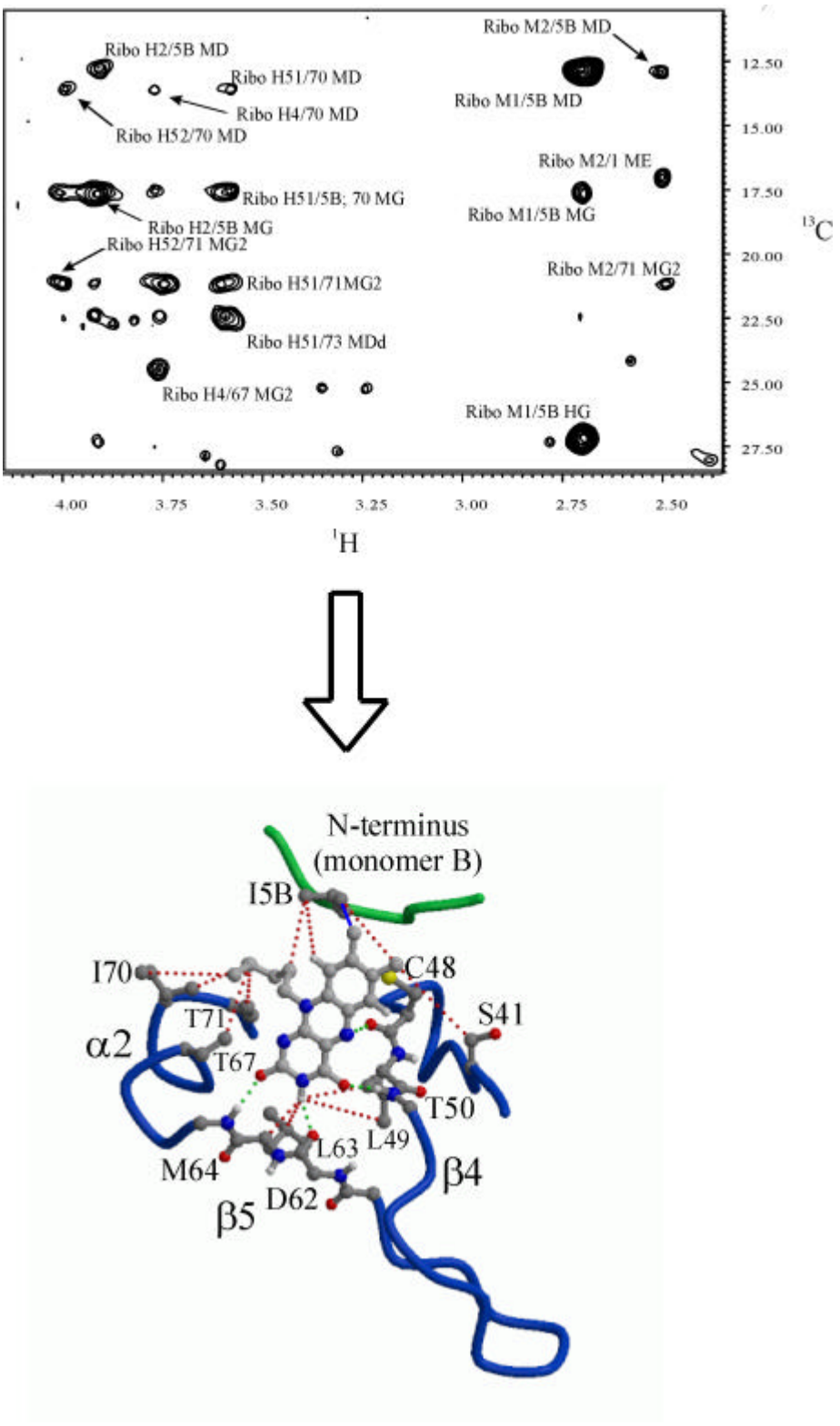


Fig. 3.22: *The $^{12}C/^{14}N$-filtered 2D-$^{13}C(H)$-NOESY spectrum (upper panel) shows contacts between both monomers of RiSy-N and the riboflavin. It allows a very precise definition of the r iboflavin orientation in its binding pocket. Monomers A and B are shown as blue and green traces, respectively (lower panel). The side-chains of residues involved in inter-molecular NOE contacts (red lines) are shown as well as expected inter-molecular hydrogen bonds (green lines).*

As explained in the *structural data* part of this chapter, ligand titration experiments previously established the binding region of the riboflavin. This result can be observed after projection of chemical shift changes on the tertiary structure. Backbone shifts in the binding region are among the most strongly perturbed by the addition of the ligand.

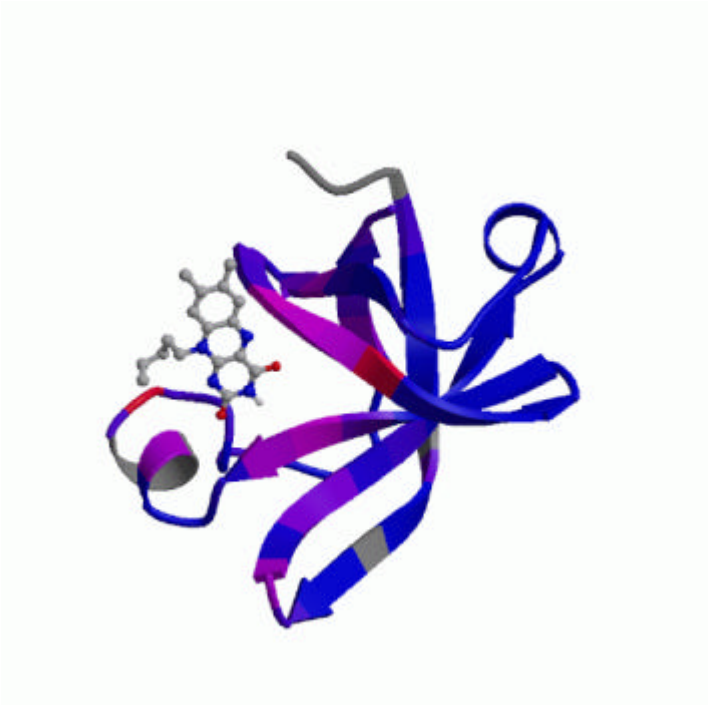


Fig. 3.23: *Coarse mapping of the chemical shift perturbation. Color coding (from blue to red) is according to the chemical shift differences between the holo- and the apo-forms of the RiSy-N monomer. Grey color stands for the HN-correlations that can't be assigned in the apo-form of RiSy-N.*

Two hydrogen bonds were also used in calculations. These were inferred from preliminary calculations and supported by structural comparison to flavin-binding proteins, as discussed in chapter 3.6.2, and involve backbone atoms of D62 and M64 (Fig. 3.22). The position of the riboflavin ring is very well-defined in the final structure set. e.g., after superimposition over one monomer (using backbone atoms within secondary structure, see Tab. 3.1), the RMSD for riboflavin ring heavy atoms is 0.37 Å. The ribose chain is not as well defined (RMSD 0.88 Å over all heavy atoms) and appears to be considerably flexible.

3.6.3 Relevance of the structure

RiSy-N and the functional trimer

A central question in the analysis of the current structure is the relevance of the dimerisation of RiSy-N to the oligomerisation of the RiSy trimer. It is possible that the dimerisation is purely artificial, i.e., it does not represent a valid model of an interface within the trimer, although this appears unlikely given its contribution to building a competent, native-like binding site. Given that the RiSy-N and RiSy-C domains share very similar structures [10, 160], it seems more likely that the dimerisation closely resembles the interaction between the RiSy-N and RiSy-C domains in the RiSy monomer. Several pieces of evidence support this, particularly the pattern of conserved residues across riboflavin synthase sequences (Fig. 3.3). The dimerisation site of RiSy-N (Fig. 3.20) is formed by a surface patch which is conserved both between and within RiSy-C and RiSy-N sequences and which includes the highly conserved residues N45/D143, N83/N181 and E85/E183 (referring to RiSy-N and RiSy-C, respectively). As these residues are also highly conserved in a related group of monomeric antenna proteins from marine bacteria (Fig. 3.3), their role in inter-domain interactions rather than in intermolecular interactions within the trimer is plausible.

Given that the dimerisation site of RiSy-N is not artificial, restraints derived from the function of the enzyme provide sufficient information to determine the organisation of the functional trimer. Plaut and Beach have observed that the regiospecificity of riboflavin synthesis (Fig. 3.2) is best explained by the formation of a triple-complex between the enzyme and two lumazine molecules which are arranged such that their heterocycles are coplanar and their ribityl chains antiparallel [183]. The presented structure confirms earlier findings of a 6:1 stoichiometry for lumazine in the functional trimer [184], i.e., each sub-domain accommodates one substrate molecule. The catalytic site must therefore be formed at an interface between sub-domains where two binding sites can be brought together. It was previously unclear whether this was an inter- or intra-molecular interface. However, the current structure excludes the intramolecular case based on

the obvious criterion that a covalent link between the RiSy-N and RiSy-C domains must be possible. Preliminary modelling studies show that it is not possible to both meet this criterion and place the bound substrates in the anti-parallel orientation required for the reaction. This immediately implies that the catalytic site is formed at an intermolecular interface. Fig. 3.20 shows that the current structure meets this criterion as the N-and C-termini of the homodimer are proximal. A plausible model of the RiSy monomer can thus be constructed simply by replacing one subunit of the homodimer by a RiSy-C structure created by homology modelling.

Given the arguments above, two arrangements of the trimer are possible. In one, each catalytic site is formed between a RiSy-N and a RiSy-C sub-domain (i.e., three NC interfaces), while the other has one NN, one CC and one NC interface. We favour the former arrangement, not only from simple arguments of symmetry, but also as two different binding sites are needed to accommodate the donor and acceptor substrate molecules [159]. The proposed model would thus allow three competent synthesis sites per trimer. We note that the proposed organisation of the functional trimer is in line with preliminary crystallographic data, which have been used to propose a pseudo-32 (D_3) symmetry [159, 160]. The three RiSy subunits are related by a threefold axis, while the two domains of each subunit are related by a pseudo-twofold axis. Each RiSy subunit therefore behaves as a pseudo-dimer, consistent with the structure presented here.

RiSy-N and the X-ray structure

As discussed above, we were able to make proposals for the structure of the whole enzyme and the organisation of the functional trimer based on the homodimer structure. Simultaneously to our work and after submission of our publication [185], the crystal structure of the functional trimer was reported by Liao et al [186], which confirmed our predictions as well as earlier ones based on sequence analysis and functional data. The N- and C-domains share very similar β-barrel structures, with each containing a substrate binding site capable of accommodating one lumazine molecule. The two domains are arranged together

such that their binding sites are on opposite faces of the protein (Fig. 3.20), so the active site must be created at an intermolecular interface. This is achieved by trimerisation, conferred by a long helical extension to the C-terminal domain. The trimer is, however, not as symmetric as expected, with only two of the six ligand binding sites arranged close enough together to form an active site (Fig. 3.24).

Fig. 3.24: *Crystal structure of the 75 kDa riboflavin synthase*[186]*. The view looks down the threefold axis. Each subunit has a (pseudo) twofold axis. The three subunits are coloured green, blue and red, with C-domains lighter than N-domains.*

The structural similarity of the RiSy β-barrel to that of the flavin binding domains of flavodoxin reductases was also noted by Liao *et al*. This led to the suggestion that the two proteins would bind their ligands in similar ways, a proposal confirmed by the inclusion of riboflavin as a bound ligand in the solution structure.

The binding site is formed by a shallow groove located at one end of the barrel and is formed partially from elements of both domains (Fig. 3.20). The structures of the apo- and holo-forms of the protein appear to be very similar, although slight differences in the crystal and solution structures exist at the N-terminus. In the solution structure the first three residues adopt an extended conformation and make contacts with the other domain. In contrast, the N-terminus of the crystal structure turns back on itself and forms an almost identical set of contacts within the N-domain (Fig. 3.24). Careful examination of our data reveals that these differences are real and chemical shift differences between the apo- and holo-forms do not indicate that they are due to ligand binding. It therefore seems likely that they arise as an artefact of homodimerisation. The binding site is almost unaffected by this difference, which is somewhat surprisingly, considering the role of the N-terminal strand for its formation, and the RiSy-N homodimer can be considered to be a very good model for ligand binding in the whole enzyme. Also, this demonstrates the quality of the solution structure of RiSy-N that can be superimposed to the crystal structure with an RMSD of 1.33 Å.

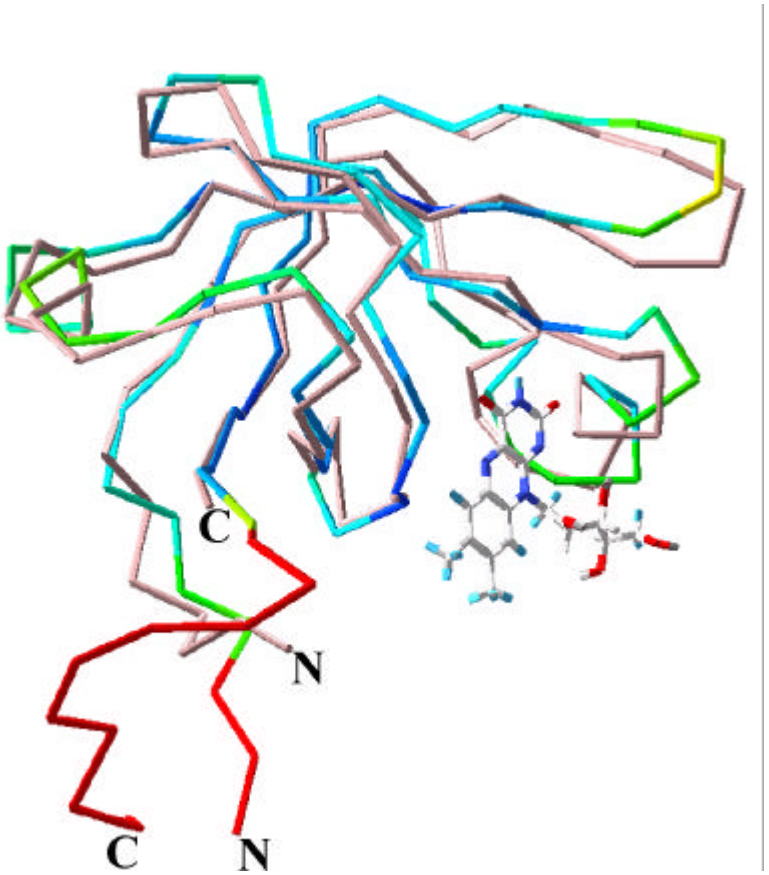


Fig. 3.25: *Superimposition over the C_α trace of the crystal structure (grey) and the solution structure (colored). Color coding (from blue to red) shows a very good alignment of the secondary elements (blue) with a backbone RMSD of 1.33Å.*

Two questions are raised by the asymmetric organization of the functional trimer. The first is the problem of a protein which uses only one third of its potential catalytic machinery. The other is that the two substrates are bound in shallow binding sites on the enzyme surface but must have both sufficient room to diffuse into the active site and simultaneously be arranged in an roughly coplanar orientation. The obvious solution to this problem is a change in the trimer structure, with one open conformation allowing free access to the binding sites and another closed conformation where the substrates are brought together. The crystal structure would then represent one possible combination of open and closed sites, with all the others also used in a cyclic manner, each transition from closed to open forming another closed site.

In many organisms riboflavin synthase is closely associated with the enzyme which catalyses the formation of its substrate, lumazine synthase [151, 152].

Lumazine synthase is a very large complex of 60 copies of a single protein arranged into a hollow capsid. The entire riboflavin synthase trimer is contained within this capsid. The efficiency of this arrangement is clear if it is considered that the second product of riboflavin synthesis, 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione, is reused by lumazine synthase. This monocyclic compound is always present at high local concentration, ready to participate in lumazine synthesis. Any proposal of cyclic conformational changes in RiSy during synthesis would imply free movement of RiSy within the lumazine synthase capsid, which analysis of the respective structures shows to be plausible. That RiSy is not regularly arranged within the caspid is evidenced by the lack of traceable electron density for RiSy in crystal structures of lumazine synthase, although whether this is due to motion of RiSy itself is not clear. The idea that RiSy functions as a kind of molecular machine is therefore a very interesting possibility.

4 Reaction Mechanism

4.1 General knowledge at the beginning of this work

Despite considerable efforts over several decades, the mechanisms of the dismutation reaction are still not completely understood. The most widely accepted mechanism for riboflavin synthesis is that of Beach and Plaut [148, 187]. It involves the formation of an exomethylene anion as an early reaction step. This idea is based on the acidity of the 7-methyl group of lumazine, the protons of which exchange with water under non-enzymatic conditions, with increased rates of exchange both at acidic and basic pH [188]. The protons of the 6-methyl group, in contrast, are non-exchangeable. Using deuterated and tritiated substrates, Plaut and Beach found an isotope effect on reaction rates only for the 6-methyl group. Deuteration at this position slowed the production of riboflavin by a factor of five, but did not effect the rate of exchange at position 7 [189]. These results were taken together to postulate a reversible deprotonation at the 7-position prior to a rate determining deprotonation at position 6 as key steps in the reaction mechanism. This is consistent with the non-enzymatic, acidic equilibrium shown in Fig. 4.1a, and the authors proposed hydrogen bonding between an amide group on the protein (labelled HX in Fig. 4.1b) to the 2-oxo group of lumazine as a mechanism for stabilising the exomethylene form [189].

Another consequence of the acidity of the 7-methyl group of lumazine is the susceptibility of the 7-position to nucleophilic attack. For example, attack by hydroxyl groups of the ribityl chain can occur non-enzymatically resulting in the formation of cyclic ethers [190]. The inhibition of the enzyme by sulfhydryl binding reagents was used to suggest that a free cysteine residue acts as a nucleophile in the reaction mechanism (labelled Z in Fig. 4.1b) [189]. Combining this data, Beach and Plaut proposed a mechanism beginning with the deprotonation of the 7-methyl group on the acceptor lumazine (performed by the proton acceptor labelled A: in Fig. 4.1b), with a simultaneous nucleophilic attack at the 7 position of the donor lumazine.

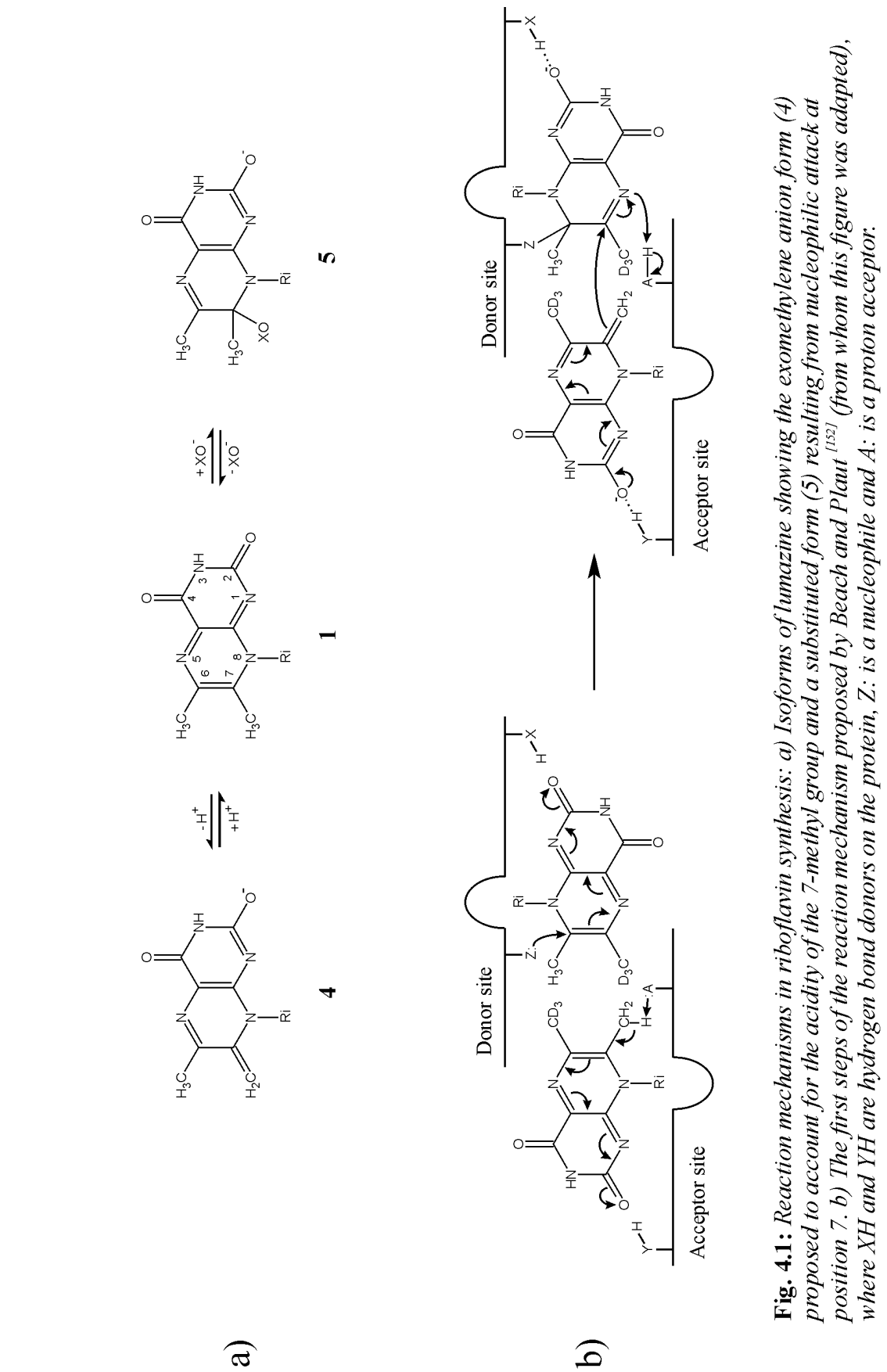


Fig. 4.1: *Reaction mechanisms in riboflavin synthesis: a) Isoforms of lumazine showing the exomethylene anion form (4) proposed to account for the acidity of the 7-methyl group and a substituted form (5) resulting from nucleophilic attack at position 7. b) The first steps of the reaction mechanism proposed by Beach and Plaut [152] (from whom this figure was adapted), where XH and YH are hydrogen bond donors on the protein, Z: is a nucleophile and A: is a proton acceptor.*

An attack on the 6 position of the donor lumazine was proposed to follow, consistent with the regioselectivity observed for the reaction. The basic requirements of the enzyme in accordance with this mechanism are shown in Fig. 4.1b.

The newly available structural data supports many of the proposals made by Beach and Plaut. For example, the 2-oxo group of the ligand is hydrogen bonded to the backbone amide of M64 in the RiSy-N binding site (Fig. 3.20). One cysteine residue, C48, is absolutely conserved in all riboflavin synthase sequences and lies within the binding site, well positioned to perform the nucleophilic role. The basic group within the acceptor site required to perform the deprotonation of the 7-methyl group could be H102, which is also completely conserved and well positioned.

However, other aspects of the reaction are not so consistent. Most of the data available on the enzymatic reaction, e.g., the relatively modest increase in reaction rate over the non-enzymatic reaction and the lack of “strong” reaction centres such as metal ions, prosthetic groups or cofactors, suggests that the role of the enzyme is predominantly entropic, with a step requiring exact alignment of the substrate molecules occuring very early in the reaction.

This state of affairs has recently been dramatically altered by the appearance of two independent structural studies, the crystal structure of the enzyme of *E. coli* [186] and the solution structure of the homodimer formed by the N-terminal domain (residues 1-97) (RiSy-N) [185], and by the isolation of a pentacyclic intermediate [191] (Fig. 4.2). The solution structure, solved in our laboratories, included riboflavin as a bound ligand and thus provides detailed information on ligand binding.

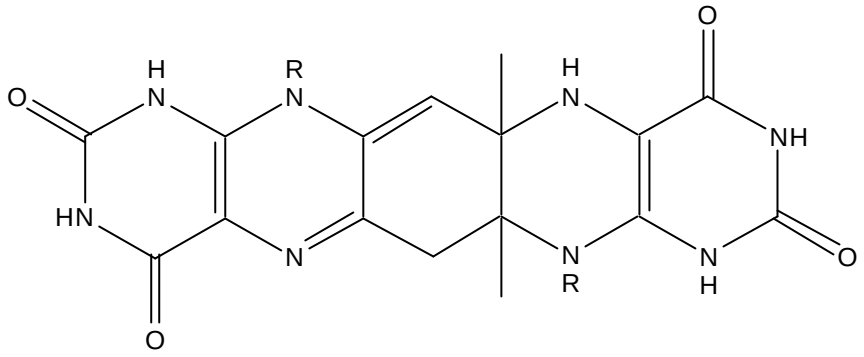


Fig. 4.2: *The pentacyclic intermediate isolated by Illarionov et al.* [191] *after a S41A point mutation.*

Taken together these three studies provide considerable new insights into the nature of the substrate binding and catalytic sites, on which the proposed reaction mechanism is based. It has led us to consider a mechanism which begins rather differently, namely with a hetero-Diels-Alder reaction with inverse electronic demand.

4.2 Riboflavin synthase a Diels-Alderase?

Diels-Alder reactions are well known reactions in organic chemistry, but are rarely found in biological systems, and instances of Diels-Alder reactions in enzymatic process have only recently been reported [192-195]. Macrophomate synthase [192] from the fungus *Macrophoma commelinae* catalyses an intramolecular Diels-Alder reaction in the formation of a bicyclic intermediate from the addition of pyruvate enolate to a 2-pyrone in macrophomic acid biosynthesis. Lovastatin nonaketide synthase [193] from *Aspergillus terreus* catalyses an intermolecular Diels-Alder reaction in the cyclization of a triene to a decalin in lovastatin biosynthesis. Although the chemistry of the two reactions is rather different, the enzymes appear to share some common features. In both cases the enzyme appears likely to assist closure via the formation of a hydrogen bond between the substrate and a carbonyl oxygen on the protein, thereby making the dienophile more electron-deficient. This essentially resembles Lewis acid catalysis of Diels-Alder addition. Also, in both cases the enzyme stabilises the substrate in its planar starting form but not the non-planar product which diffuses out of the enzyme.

The mechanism of riboflavin synthesis can also be rationalised in these terms. The fact that no catalytic residue and that only efficient interaction of π-orbitals between the two substrates are required in a Diels-Alder reaction; the high stereo-specificity that is a typical property of the Diels-Alder reaction and the absence of products that possess only one C-C bond between the substrates, provide circumstantial evidence for the Diels-Alder route mechanism. The proposed mechanism differs from other Diels-Alder enzymes in the nature of the reaction itself. We propose a hetero-Diels-Alder reaction with inverse electron demand as the first step in the dismutation. The suggested mechanism is shown in Fig. 4.3. In contrast to the other Diels-Alder enzymes, the inverse electron demand would be assisted by electron enrichment of the dienophile coupled to withdrawal of electrons from the diene. Subtle differences in the N- and C-terminal binding sites could create these conditions, therefore differentiating acceptor and donor sites.

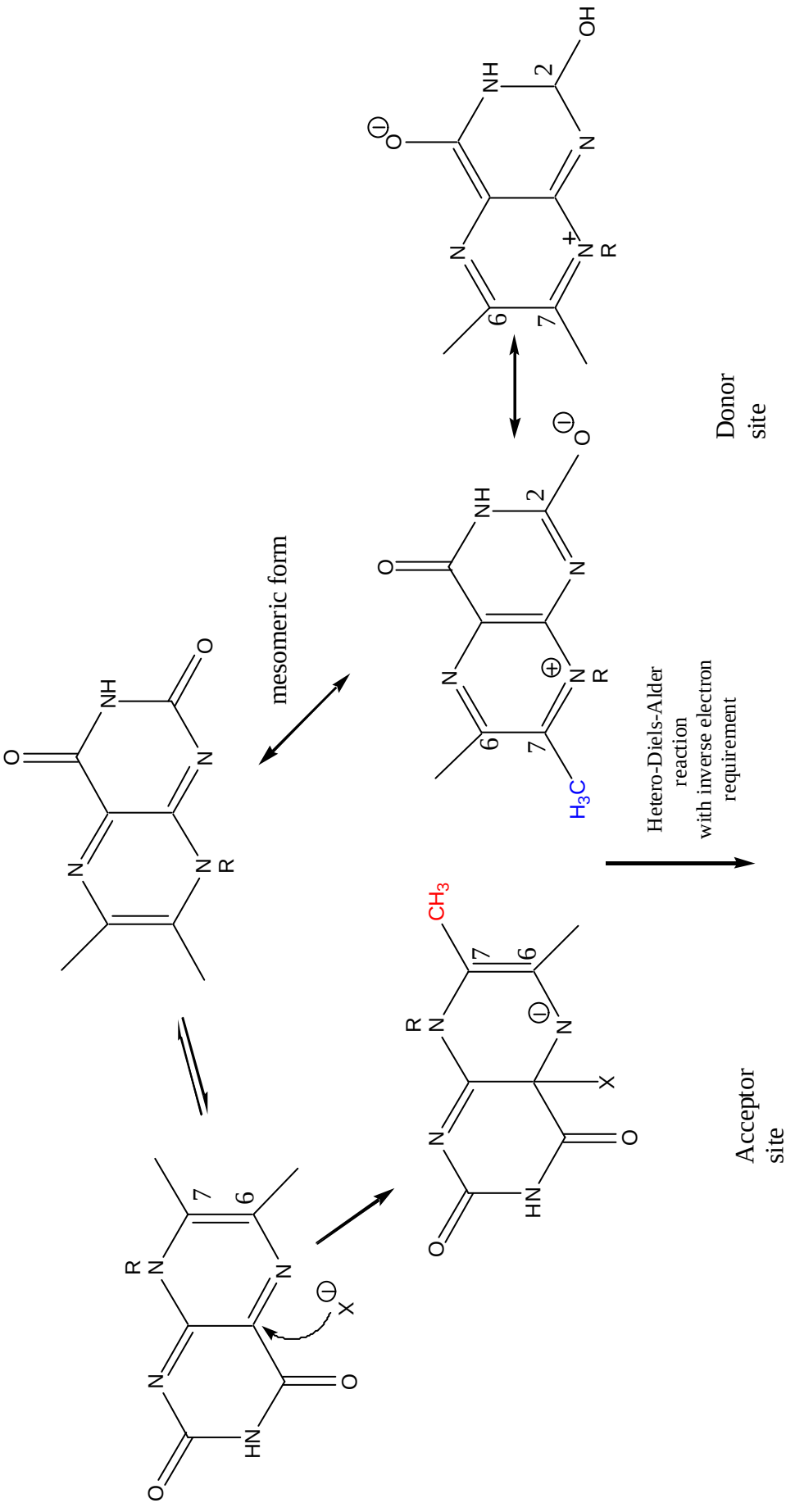
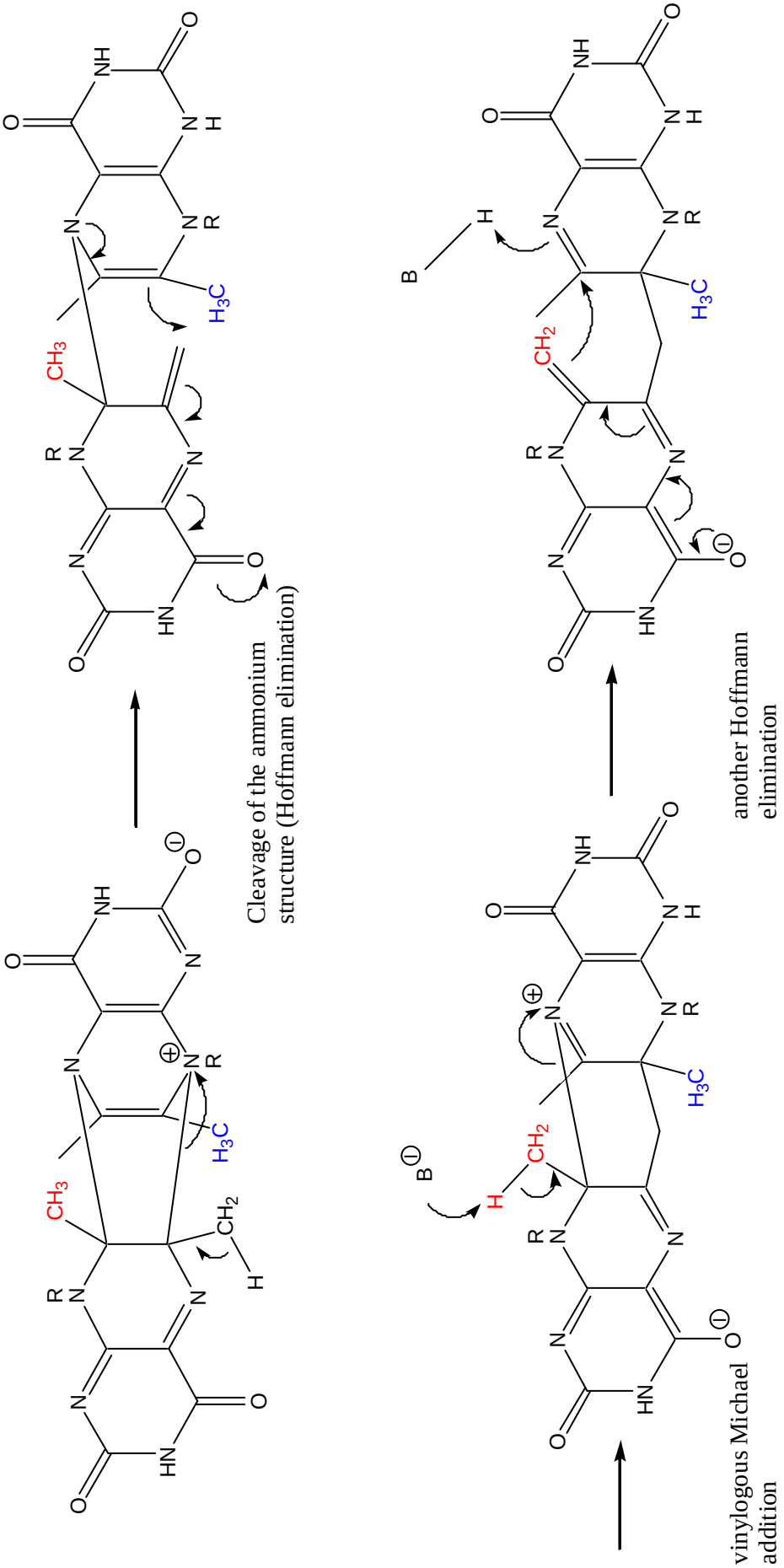


Fig. 4.3: The reaction mechanism of riboflavin synthesis via a Diels-Alder reaction as a first step. The regiospecificity of the reaction is shown as red and blue labeling of the concerned methyl groups. X represents a nucleophile and B a base.


vinylogous Michael addition
another Hoffmann elimination
Cleavage of the ammonium structure (Hoffmann elimination)

5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione
+
Riboflavin
elimination and aromatization
retro Michael addition
aldol-type addition
deprotonation (activated Hydrogen!)

Indeed, the RiSy monomer contains two distinct binding sites which should bind the donor and acceptor lumazine molecules specifically. The current study predicts that the two binding sites are very similar, one comprised of residues from β4, β5 and α2 of RiSy-N with a contribution from the N-terminal hexapeptide of RiSy-C (the N-site) and the other of β4, β5 and α2 of RISY-C with a contribution from the N-terminal hexapeptide of RiSy-N (the C-site) (Fig. 3.20). The binding site of the homodimer therefore represents a chimera of the donor and acceptor sites and may have hybrid binding properties.

These sites have, however, rather different binding properties. While lumazine and bicyclic precursor analogues are bound at both sites, riboflavin and tricyclic product analogues are bound at one site only, presumable the acceptor [184]. Further, a monocyclic analogue of the second synthesis product is also bound at only one site, which should be the donor [184]. Given the high affinity of the RiSy-N homodimer for riboflavin, it would be tempting to assign the acceptor role to the N-site.

In the donor site, the hydrogen-bonds between the backbone NH of I162 and the 2-oxo group of lumazine and between T147 backbone NH and the 4-oxo group of riboflavin together stabilize the lumazine in the charged form shown in Fig. 4.1a, resulting in an electronically depleted diene. Although M64 and T50 are able to play the same role in the acceptor site, the difference comes from the well conserved C48. Both ourselves and Liao *et al*. have noted the possible formation of a protease-like nucleophilic triad involving C48, H102 and E183. A model of the RiSy-N binding site can be made via homology, replacing residues from the N-terminal strand with the equivalent residues from RiSy-C, i.e., residues L98-I103 (data not shown). It is then possible that the histidine accepts a positive charge to activate C48 as a nucleophile. A third residue with a similar conservation pattern, E183 from RiSy-C, is well positioned to represent the acidic residue whose negative charge stabilises the charged histidine, thus completing the triad.

Our structural studies show that C48 is perfectly positioned to perform a nucleophilic attack at the carbon 10 of the lumazine, thus giving an electron rich

form of lumazine stabilized by the hydrogen bond between T50 and the 4-oxo group of riboflavin (Fig. 4.1b). The second step involves a basic group susceptible to react with the 6-methyl group. These protons were found to be non-exchangeable [188], which means that this reaction needs some specific conditions to act. We discussed above the presence of C48 as a nucleophilic residue. Although C48 can't be considered as a typical basic residue, it is basic enough to assist the deprotonation of the 6-methyl group nearby the positively charged quaternary amide. On the last determinant step of the reaction, a proton from the exchangeable 7-methyl group in the acceptor site will be transfered to the nitrogen-5 at the donor site. This deprotonation role could be achieved by H102. This histidine is well placed in the binding site, close to both the methyl-7 and the nitrogen-5, and we suggest that the protonated histidine will give rise to a "flip-rotation" to give the proton back to the nitrogene-5 and thus forming a pentacyclic intermediate via an aldol-type addition. The last steps of the reaction mechanism involve well-known organic reactions which do not require any particular assistance from the enzyme. The reaction ends with a retro Michael addition giving the products after elimination and aromatisation. The suggested mechanism satisfies the regiospecificity of the reaction elucidated by Schott *et.al.* [159] in accordance with the regiospecificity expected for a Diels-Alder reaction (Fig. 4.2).

An important intermediate of the reaction is the pentacyclic form which was isolated, via a S41A mutation, in the group of Prof. Bacher [191]. Although S41 is well conserved, its function is not easily determined, as it is somewhat distant from the active site. Another well conserved residue, F2, is close to S41. A structural change could be envisaged resulting by π-stacking interactions between the phenylalanine and one of the aromatic cycles of the intermediate, stabilising the pentacyclic compound in the active site, but the exact role of both of these residues is unclear. However, the transition state is highly ordered and very closely resembles the products, as expected for a Diels-Alder mechanism.

An important point concerns the capacity of the enzyme to stabilize both the substrates at the beginning of the reaction and the pentacyclic intermediate. As

described for macrophomate synthase [192], we propose that the multicyclic intermediate is essentially stabilised by the groups used for stabilisation of the substrates.

An interesting difference between MPS, LNKS and RiSy is the size of the proteins themselves. The fact that the active site in the riboflavin synthase is formed at the interface of two domains permits more radical structural changes and thus the stabilisation of both the coplanar substrates at the beginning of the reaction and the non-planar pentacyclic intermediate later. This is not possible for the relatively small MPS and LNKS proteins, where the intermediates diffuse out the protein immediately after the Diels-Alder reaction. The observation of the pentacyclic intermediate proves the involvement of the protein till this step.

This idea is also consistent with the asymmetry of the crystal structure that can thus explain the cyclic conformational changes within the functional trimer discussed earlier. RiSy can then be considered as a machinery where an open active site is free of ligand, waiting for a new binding of substrates. Simultaneously, a second active site is reacting as the last one releases the products. In case of the complex RiSy/LuSy from *Bacillus subtillis* [151, 152], the by-product will be easily recycled thus giving to this machinery its high efficiency.

5 Summary

In the present work, the solution structure of the N-terminal domain of riboflavin synthase was investigated.

In chapter 2, structure determination by NMR was discussed in detail. Collection and analysis of information, necessary for the elucidation of protein structures, was presented step by step, also offering several approaches that allow the study of large biological macromolecules. Recent methodological developments that are actually used for the investigation of large protein complexes up to 65 kDa were pointed out. Nevertheless, the assignment process of backbone and side-chain resonances is a very long and laborious manual work. Also several computer programs are available to alleviate this step. In particular, PASTA [66] allowed a semi-automatic assignment which results in 99 % backbone and 94 % side-chain shifts assigned for RiSy-N.

The solution structure of the N-terminal domain of riboflavin synthase presented in chapter 3 was solved with the aid of NMR spectroscopy. The lack of significant sequence similarity to proteins of known structure had previously precluded homology-based structure predictions. However, riboflavin synthase forms part of a protein family together with lumazine antenna proteins and yellow fluorescent protein. The structure displays a six stranded β-barrel, typical of a greek-key scheme where both ends of the barrel are closed by an α-helix. The structure and the binding pocket of RiSy-N are very similar to those of flavodoxin reductase [174] and related flavin-binding proteins [175-177]. RiSy-N behaves as a homodimer and shows a very high degree of symmetry up to the repetition of two ββαβ-motives along its primary sequence. In order to overcome the difficulties in differentiating between intra- and inter-NOE contacts, methodological developments were established in our laboratories. They allow the precise description of the site of dimerisation of RiSy-N and of the binding mode of the riboflavin. The development of a new isotope filter experiment was especially helpful in the latter

case where 11 ligand-protein contacts were observed. Taken together, these findings allow us to propose the organization of the functional trimer. The homotrimeric protein possesses a pseudo-D_3 symmetry [160] with three identical active sites. Each active site is formed at the interface of two subunits by a C- and an N-terminal domain. Each domain binds one lumazine molecule, which can react with each other once brought together.

Finally, the reaction mechanism was investigated, based on insights provided by the NMR- and the X-ray structures. The first step of the proposed mechanism involves a hetero-Diels-Alder reaction with inverse electron demand. A nuceophilic attack by the free cysteine C48 on the acceptor lumazine and some specific hydrogen bonds between the enzyme and the donor lumazine are implicated in catalysis. In the last key step of the reaction, a proton from the exchangeable 7-methyl group [188] in the acceptor site will be transfered to the nitrogen-5 at the donor site. The role of the deprotonating agent might be played by H102. Also, we note the possible formation of a protease-like nucleophilic triad involving C48, H102 and E183. The fact that no "strong" reaction centres such as metal ions, prosthetic groups or cofactors are present and that only efficient interaction of π-orbitals between the two substrates is required in a Diels-Alder reaction, provide circumstantial evidence for the Diels-Alder route mechanism. It is suggested that the enzyme assists the reaction up to the formation of the kinetic pentacyclic intermediate [191]. The mechanism also satisfies the regiospecificity of the reaction, elucidated by Schott *et.al.* [159].

After all, riboflavin synthase appears to be a molecular machinery where each active site successively gives rise to the enzymatic reaction. Experiments, on the base of molecular modelling, are still on going in our laboratories, trying to verify these proposals.

6 Appendix

A Spectrometer specifications

Spectrometer	Probehead	Gradients calibration [a]	Calculated T_{real} from $T_{indicated}$ [K]	0 ppm [^{1}H; MHz]	0 ppm [^{13}C; MHz]	0 ppm [^{15}N; MHz]
DMX750	TXI	(74%)	$0.996123 \cdot T_{indicated}$ - 0.624249 K	750.1299379	188.6198203	76.0100050
DMX600	QXI	66 ± 0.7 (74%)	$1.031109 \cdot T_{indicated}$ - 11.942032 K	600.3999475	150.9702846	60.8379971
DMX601	TXI	67.0 ± 0.5 (71.5%)	$0.9476 \cdot T_{indicated}$ + 13.055	600.1299400	150.9023900	60.8106370
DMX500	BB-inverse		$1.0352 \cdot T_{indicated}$ - 5.3761	500.1299577	125.7574459	50.6777295

[a]The first entry renders the absolute magnitude of z -gradients in [G/cm] corresponding to 100% amplitude. The second entry (in brackets) refers to the ratio g in absolute magnitude of x,y-gradients with respect to the z-gradient given they all experience the same nominal gradient strength. In order to achieve magic angle gradient pulses the z-coil nominal amplitude must be scaled down by g. TXI/QXI and BB-inverse probes are equipped with xyz- and z- gradients coils, respectively.

B Data appendix of RiSy-N

B 1: RiSy-N spectral regions

SIGNAL	WINDOW/ RANGE [ppm]	SFO (DMX600) [MHz]	SFO (DMX601) [MHz]	SFO (DMX750) [MHz]
Protons	-2.17 - 11.78	600.4028262 increment: 59.6µs	600.1328250 increment: 59.63µs	750.1335238 increment: 47.6µs
$H_{aliphatic}$	-0.47 – 5.75	150.9774557 increment: 270µs	150.9095611 increment: 270.12µs	188.6287722 increment: 216.11µs
H_{α}	2.99 - 5.75	-	-	-
H_{β}	0.25 - 4.59	-	-	-
$H_{aromatic}$	6.46 – 9.96	-	-	-
$H^{N,backbone}$	6.92 - 11.06	-	-	-
$H^{N,sidechain}$	5.82 - 8.23	-	-	-
Carbons				
$C_{aliphatic}$	8.30 - 81.80	150.9770783 increment: 90µs	150.9091839 increment: 90.04µs	188.628310 increment: 72µs
$C_{aliphatic,folded}$	48.55 - 85.50	150.9803996 increment: 179µs	150.9125037 increment: 179.02µs	188.6326063 increment: 143.28µs
C_{α}	44.34 - 66.04	-	-	-
C_{β}	27.09 - 72.92	-	-	-
$C_{aromatic}$	114.11 – 128.63	150.9895330 increment: 190µs	150.92163301 increment: 190.08µs	188.6438614 increment: 152.08µs
CO	172.12 - 184.61	150.9973083 increment: 470µs	150.9294048 increment: 470.22µs	188.6537317 increment: 376.18µs
Nitrogen				
$N_{backbone}$	106.12 – 132.30	60.8452490 increment: 627.3µs	60.8178855 increment: 627.58µs	76.019060 increment: 502µs
$N_{Asn\text{-}N}{}^{\delta}$	107.03 – 113.88	-	-	-

B 2: RiSy-N sample conditions

RiSy-N Sequence: 97 amino acids

MFTGI VQGTA KLVSI DEKPN FRTHV

VELPD HMLDG LETGA SVAHN GCCLT

VTEIN GNHVS FDLMK ETLRI TNLGD

LKVGD WVNVE RAAKF SDEIG GH

Protein Properties:

Melting Point. 50°C (323 K)

Experimental conditions:

Temperature: 300 ± 0.5 K

Solvent: H_2O/D_2O: 9/1

pH: 7.3

Concentration: 1 mM

Minimal medium and inoculum cultures

The minimal medium was derived from M9 minimal medium (Maniatis) and had the following composition: 62 mM Na_2HPO_4, 65 mM KH_2PO_4, 62 mM K_2HPO_4. The medium was autoklaved, chilled to room temperature and supplemented with 25 µM $CaCl_2$, 2 mM $MgSO_4$, trace elements from a 1000 fold stock solution (68 mM $MnCl_2$, 30 mM $FeSO_4$, 3 mM $ZnSO_4$, 1.6 mM H_3BO_3, 1.1 mM $CuCl_2$, 0.8 mM $NiCl_2$, 0.4 mM $CoCl_2$ and 170 mM EDTA adjusted to pH 7.0 prior to sterilisation), Vitamins from a 250 fold stock solution (20 mg Pyridoxamin-HCl, 40 mg Thiamin-HCl, 20 mg Riboflavin , 20 mg Ca-Pantothenat, 20 mg Biotin, 10 mg Folsäure, 0.1 mg mg Cyanocobalamin per liter stock solution). A feed

solution a with 2,77 M Glucose, 1.86 M NH_4Cl, 4mM $MgSO_4$ was used for fed batch fermentation.

For inoculum cultures 0,2 % (w/v) glucose and 1 g/L NH_4Cl were added to the medium. 100 mL of this medium was inoculated (1:200) with a overnight culture grown on Luria broth and incubated at 37°C for 14 h.

Fermentation

The fermentation was started as a batch fermentation in a 1,25 L vessel (Bioflo3000, Newbrunswick Scientific). The batch medium had the same composition as the medium used for inoculation cultures. The pH was kept at 6.8 by the addition of 1M NaOH. The medium (1 L) was inoculated with 50 mL inoculation culture. After oxygen saturation dropped down to 30% it was held constant by a DO-agitation-loop regulation (controller feature of Bioflo 3000). After depletion of glucose, marked by a sharp increase in oxygen saturation, a feedback controlled fed batch was started by using the sharp increase in oxygen saturation for triggering a pump with the feed solution. The consumption of NH_4^+ was monitored by a ammonium-test kit purchased from Merck Kg (Darmstadt, Germany) and supplemented manually if needed with additional NH_4Cl from a 200g/L stock solution. The fed batch culture was grown to a OD_{600} of 5. At this point it was induced with 5 mM IPTG and grown until the feed solution was completely consumed (OD_{600}= 11 for a total amount of 6 g glucose). The cells were harvested by centrifugation, washed once in 0.8% NaCl and stored at –70°C.

The complete expression and purification of the recombinant N-terminal domain of riboflavin synthase of *E. coli* was described by Eberhardt *et al.* [161] (and references therein).

B 3: NMR experiments that were carried out for the structure determination of RiSy-N

Experiments for backbone assignment

EXPERIMENT	SPECTROMETER	COMMENTS
^{15}N-HSQC	DMX 600	high resolution: reference spectrum
^{15}N-HSQC	DMX 600	high resolution: reference spectrum [106.12-132.30; ppm] (folded)
^{13}C-HSQC	DMX 600	high resolution: reference spectrum
^{13}C-HSQC	DMX 600	high resolution: reference spectrum [48.55-85.50; ppm] (folded)
HNCA	DMX 600	$\Delta(NC_\alpha)$ = 14 ms and sensitivity enhancement
HNCACB	DMX 600	$\Delta(NC_\alpha)$ = 12.6 ms, $\Delta(C_\alpha C_\beta)$ = 3 ms
CBCA(CO)NH	DMX 600	$\Delta(NCO)$ = 12 ms, $\Delta(C_\alpha CO)$ = 3.6 ms, $\Delta(C_\alpha C_\beta)$ = 3.25 ms
HBHA(CO)NH	DMX 600	$\Delta(NCO)$ = 13 ms, $\Delta(C_\alpha CO)$ = 3.6 ms, $\Delta(C_\alpha C_\beta)$ = 3.25 ms
HN(CA)HA	DMX 601	$\Delta(NC_\alpha)$ = 12.6, $\Delta(C_\alpha H_\alpha)$ = 1.5 ms
HNHA	DMX 600	$\Delta(NH_\alpha)$ = 13 ms
HNHB	DMX 601	$\Delta(NH_\beta)$ = 26.5 ms, simultaneously rephrasing $^1J_{HN}$
HNCO	DMX 600	$\Delta(NCO)$ = 13 ms
HN(CA)CO	DMX 600	$\Delta(NC_\alpha)$ = 14 ms, $\Delta(C_\alpha CO)$ = 3.6 ms

Experiments for side-chains assignment

EXPERIMENT	SPECTROMETER	Comments
(H)CCH-COSY	DMX 600	$\Delta(C_1C_2)$ = 3.28 ms, $^{13}C_{aliphatic}$ [48.55-85.50; ppm] (folded)
H(C)CH-TOCSY	DMX 600	τ_{mix} = 19.5 ms, $^{13}C_{aliphatic}$ [48.55-85.50; ppm] (folded)

Experiments for NOE assignment

EXPERIMENT	SPECTROMETER	COMMENTS
HNH-NOESY	DMX 600 DMX 750	τ_{mix} = 80 ms, old sample τ_{mix} = 80 ms, new sample
HCH-NOESY	DMX 750	τ_{mix} = 80 ms, ^{13}C: [48.55-85.50; ppm] (folded)
CNH-NOESY	DMX 750	τ_{mix} = 80 ms
NCH-NOESY	DMX 750	τ_{mix} = 80 ms
CCH-NOESY	DMX 750	τ_{mix} = 80 ms
NNH-NOESY	DMX 601 DMX 750	τ_{mix} = 80 ms, old sample τ_{mix} = 80 ms, new sample

Experiments for determination of binding and dimerisation sites

EXPERIMENT	SPECTROMETER	COMMENTS
^{15}N{^{1}H}-NOE	DMX 750	pseudo 3D, fully interleaved Heteronuclear NH-NOE 3 s ^{1}H presaturation
MEXICO	DMX 600	pseudo 3D, water exchange experiment τ_{mix} = 50, 100, 150, 200, 250, 300 ms
HCH-NOESY	DMX 600	^{13}C-edited 2D-NOESY on ^{15}N-RiSy-N + ^{15}N/^{13}C-riboflavin τ_{mix} = 100 ms
HCH-NOESY	DMX 600	^{12}C/^{14}N-filtered 2D-^{13}C(H)-NOESY on ^{15}N/^{13}C-RiSy-N + unlabelled riboflavin, τ_{mix} = 100ms
^{15}N-HSQC	DMX 600	riboflavin free, [U-^{15}N]-RiSy-N
^{15}N-HSQC	DMX 600	10% riboflavin, [U-^{15}N]-RiSy-N + [U-^{13}C, ^{15}N]-riboflavin
^{15}N-HSQC	DMX 600	33% riboflavin, [U-^{15}N]-RiSy-N + [U-^{13}C, ^{15}N]-riboflavin
^{15}N-HSQC	DMX 600	66% riboflavin, [U-^{15}N]-RiSy-N + [U-^{13}C, ^{15}N]-riboflavin
^{15}N-HSQC	DMX 600	100% riboflavin, [U-^{15}N]-RiSy-N + [U-^{13}C, ^{15}N]-riboflavin
^{13}C-HSQC	DMX 600	100% riboflavin, [U-^{15}N]-RiSy-N + [U-^{13}C, ^{15}N]-riboflavin

Experiments for determination of the oligomerization state

EXPERIMENT	SPECTROMETER	COMMENTS
CNH-NOESY	DMX 600	cross-labeled sample, 128 scans, τ_{mix} = 80ms
2D-CNH-NOESY	DMX 750	cross-labeled sample, 1024 scans, τ_{mix} = 80ms

Diffusion experiments

EXPERIMENT	SPECTROMETER	COMMENTS
Diffusion	DMX 601	water diffusion, gradient ramp [2-40%], τ_{mix} = 28ms
Diffusion	DMX 601	Protein diffusion, gradient ramp [2-80%], τ_{mix} = 128.7ms

-Δ(XY) = half length of the scalar coupling evolution between spins X and Y, in reference to the INEPT-transfer (Δ–180°–Δ).
- τ_{mix} = mixing time for the population transfer (NOE-transfer) or coherence transfer (TOCSY-transfer)

B 4: Experimental 15N{1H}-NOE and water exchange data for RiSy-N

Amino acid	^{15}N [ppm]	^{1}H [ppm]	Secondary structure	H_2O K_{ex} [Hz]	NOE 750 MHz
M1					
F2	116.85	8.94		2.24	0.872885
T3	115.83	11.06		2.22	0.908568
G4	111.53	9.61		3.14	0.849071
I5	122.59	8.14			
V6	126.05	8.56			0.953604
Q7	124.84	7.96		2.27	0.731975
G8	106.61	7.81	β1	1.81	0.826811
T9	108.46	8.24	β1	1.63	0.751134
A10	121.56	9.18	β1	1.89	0.900702
K11	124.83	7.73	β1	1.82	0.803713
L12	127.29	8.68	β1	1.54	0.846476
V13	124.22	8.92	β1	1.74	0.813631
S14	112.55	7.50	β1	1.73	0.888072
I15	122.38	8.20	β1	1.63	0.791554
D16	126.88	8.95	β1	1.59	0.818878
E17	125.65	8.86	β1	3.14	0.757415
K18	127.49	8.29	β1	1.65	0.734161
P19					
N20	110.60	8.28		1.67	0.741451
F21	115.21	7.29	β2	2.13	0.869589
R22	118.29	8.77	β2	1.58	0.755972
T23	122.79	9.34	β2	2.85	0.732838
H24	126.68	9.75	β2	1.85	0.86249
V25	122.68	8.43	β2	1.89	0.899647
V26	118.49	9.09	β2	1.78	0.874777
E27	124.02	8.72		1.99	1.03003
L28	127.50	8.58		1.64	0.855232
P29					
D30	121.15	8.72		2.83	0.786226
H31	114.40	8.19	α1	3.19	0.817805
M32	117.67	7.35	α1	1.67	0.796745
L33	117.26	7.24	α1	1.88	0.810479
D34	119.31	7.50	α1	3.58	0.790181
G35	116.74	9.19	α1	6.12	
L36	123.29	7.69	α1	2.07	0.828615
E37	123.61	7.29		1.88	0.780015
T38	114.40	8.66		2.16	0.82299
G39	112.85	8.88		1.77	0.88603
A40	123.20	7.77		5.41	0.823865
S41	116.03	8.83	β3	1.94	0.879144
V42	123.40	9.32	β3	1.58	0.880562
A43	130.57	8.90	β3	1.97	0.791008
H44	122.59	8.59	β3	2.95	0.929372
N45	123.29	10.26			1.01364
G46	127.91	8.79		2.15	0.963803

C47	118.08	7.52		1.61	0.870956
C48	131.59	10.29			0.966945
L49	129.13	8.11			0.891199
T50	115.83	8.36			0.865404
V51	128.11	9.43	β4	2.53	0.844311
T52	121.76	8.97	β4	2.21	0.821787
E53	120.33	6.96	β4	2.04	0.670908
I54	123.4	8.66	β4	1.85	0.837077
N55	126.47	8.73	β4	1.57	0.865498
G56	116.65	9.14		1.59	0.866559
N57	123.30	8.34		1.82	0.801791
H58	119.51	8.28	β5	1.36	0.857571
V59	128.52	9.53	β5	1.55	0.907945
S60	124.02	8.69	β5	1.69	0.866434
F61	118.08	8.60	β5	1.54	0.878675
D62	123.81	10.03	β5		0.888139
L63	125.66	9.81	β5		0.928097
M64	119.72	8.64	β5	2.56	0.829733
K65	119.92	8.45		3.33	0.850243
E66	118.79	9.08			0.874777
T67	116.74	7.31			0.784607
L68	122.18	7.88	α2	1.86	0.929639
R69	116.85	7.73	α2	1.79	0.808423
I70	114.60	7.66	α2	1.93	0.760406
T71	131.28	7.16	α2	1.75	0.80738
N72	116.65	9.27		2.11	0.928006
L73	117.26	8.07		1.68	0.881103
G74	130.57	8.53		1.45	0.746666
D75	117.67	7.36		1.67	0.796745
L76	119.31	6.92		1.73	0.820501
K77	120.13	8.80		1.60	0.877756
V78	120.21	8.31			0.376151
G79	116.50	9.04		1.53	0.80361
D80	120.74	7.98		1.24	0.814834
W81	117.46	8.59	β6	1.10	0.804648
V82	118.70	9.31	β6	1.47	0.74355
N83	123.60	8.14	β6	1.40	0.864428
V84	115.01	8.68	β6		0.835847
E85	120.33	7.36	β6	1.95	0.906482
R86	127.80	9.04			0.822198
A87	123.40	7.92		3.11	-0.875031
A88	124.84	8.42		2.86	0.720924
K89	121.36	8.37			0.597655
F90	121.97	8.36		3.84	0.454451
S91	116.85	8.24		7.50	0.33261
D92	122.17	8.30		6.88	0.303348
E93	120.33	8.29		2.69	0.376151
I94	121.45	8.14		5.53	0.0583577
G95	112.24	8.44		8.24	-0.153086
G96	108.76	8.23			-0.441014
H97	123.03	7.89			

B 5: Chemical shift pertubation: assignment of holo- and apo-forms of RiSy-N

Amino acid	holo-RiSy-N ^{15}N [ppm]	^{1}H [ppm]	Secondary structure	apo-RiSy-N ^{15}N [ppm]	^{1}H [ppm]
M1					
F2	116.85	8.94		116.86	8.88
T3	115.83	11.06			
G4	111.53	9.61			
I5	122.59	8.14			
V6	126.05	8.56		125.24	8.33
Q7	124.84	7.96			
G8	106.61	7.81	β1	107.62	8.20
T9	108.46	8.24	β1	108.15	8.24
A10	121.56	9.18	β1		
K11	124.83	7.73	β1	125.14	7.75
L12	127.29	8.68	β1	127.29	8.71
V13	124.22	8.92	β1	125.03	9.00
S14	112.55	7.50	β1	112.34	7.55
I15	122.38	8.20	β1	122.37	8.24
D16	126.88	8.95	β1	126.77	8.94
E17	125.65	8.86	β1	125.75	8.85
K18	127.49	8.29	β1	127.49	8.24
P19					
N20	110.60	8.28		110.6	7.84
F21	115.21	7.29	β2	115.51	7.33
R22	118.29	8.77	β2	117.16	8.83
T23	122.79	9.34	β2	122.00	9.28
H24	126.68	9.75	β2	126.37	9.61
V25	122.68	8.43	β2	123.29	8.50
V26	118.49	9.09	β2	118.89	9.13
E27	124.02	8.72		124.21	8.80
L28	127.50	8.58		127.71	8.63
P29					
D30	121.15	8.72		120.94	8.74
H31	114.40	8.19	α1	114.40	8.20
M32	117.67	7.35	α1	117.66	7.35
L33	117.26	7.24	α1	117.15	7.27
D34	119.31	7.50	α1	119.41	7.46
G35	116.74	9.19	α1		
L36	123.29	7.69	α1	123.29	7.68
E37	123.61	7.29		123.50	7.31
T38	114.40	8.66		114.40	8.66
G39	112.85	8.88		113.06	8.98
A40	123.20	7.77		123.20	7.79
S41	116.03	8.83	β3	116.33	8.66
V42	123.40	9.32	β3	122.68	9.26
A43	130.57	8.90	β3	130.36	9.17
H44	122.59	8.59	β3		
N45	123.29	10.26		122.68	10.30
G46	127.91	8.79		128.31	8.95
C47	118.08	7.52		118.89	7.66
C48	131.59	10.29		130.97	10.15
L49	129.13	8.11			
T50	115.83	8.36		114.60	8.5
V51	128.11	9.43	β4	125.96	9.09

T52	121.76	8.97	β4	122.27	9.00
E53	120.33	6.96	β4	120.22	7.03
I54	123.40	8.66	β4	123.30	8.68
N55	126.47	8.73	β4	126.57	8.74
G56	116.65	9.14			
N57	123.30	8.34		123.30	8.34
H58	119.51	8.28	β5	119.70	8.27
V59	128.52	9.53	β5	128.52	9.57
S60	124.02	8.69	β5	124.02	8.78
F61	118.08	8.60	β5	118.79	8.71
D62	123.81	10.03	β5	122.48	10.19
L63	125.66	9.81	β5	125.44	9.63
M64	119.72	8.64	β5	120.33	8.62
K65	119.92	8.45			
E66	118.79	9.08		118.17	9.10
T67	116.74	7.31		115.51	7.13
L68	122.18	7.88	α2		
R69	116.85	7.73	α2		
I70	114.60	7.66	α2	114.39	7.48
T71	131.28	7.16	α2	106.51	7.08
N72	116.65	9.27		116.54	9.34
L73	117.26	8.07		117.97	8.20
G74	130.57	8.53		130.60	8.66
D75	117.67	7.36		118.17	7.36
L76	119.31	6.92		119.11	6.81
K77	120.13	8.80		120.63	8.85
V78	120.21	8.31		120.43	8.34
G79	116.50	9.04		116.54	9.06
D80	120.74	7.98		120.84	8.06
W81	117.46	8.59	β6	117.35	8.83
V82	118.70	9.31	β6	119.51	9.41
N83	123.60	8.14	β6		
V84	115.01	8.68	β6	115.20	8.75
E85	120.33	7.36	β6	120.02	7.51
R86	127.80	9.04		128.62	9.18
A87	123.40	7.92		123.59	7.88
A88	124.84	8.42		125.03	8.41
K89	121.36	8.37		121.36	8.54
F90	121.97	8.36			
S91	116.85	8.24			
D92	122.17	8.30			
E93	120.33	8.29		120.33	8.30
I94	121.45	8.14		121.45	8.15
G95	112.24	8.44		112.24	8.45
G96	108.76	8.23		108.55	8.21
H97	123.03	7.89			

C Pulse Programs

Most of the pulse programs used in this project are well known standard programs that were implemented from the literature. Therefore they won't be listed here but are accessible in the archives of the laboratory (directory /akkn1/risy/nmr). Only the pulse programs that were modified and optimised during the course of this thesis, e.g., 2D ^{13}C,H[^{12}C,^{14}N]-NOESY, are shown below. All these pulse programs are written for digital AVANCE spectrometers of the BRUKER company controlled by the software XWINNMR 2.6.

Modified and optimised pulse programs

Pseudo 3D ^{15}N{^{1}H}-NOE

```
;## tdnhnoese.3e ##                    td 06.11.97
;avance-version

;interleaved pseudo-3D-15N N(H) Steady-State NOE Experiment measurement
;interleaving scheme:
;1.      NOE(first scan, first t2-point)-echo
;       no-NOE(first scan, first t2-point)-echo
;       NOE(second scan, first t2-point)-echo
;       no-NOE(second scan, first t2-point)-echo
;       etc. until accumulation of desired number of scans
;2.      the same for anti-echo
;3.      increment t2 time, recommence at 1.

;FEATURES:
;using INEPT-transfers
;sensitivity enhanced
;using coherence selection via gradients
;echo/antiecho detection in F2
;Flip-back version (H2O (at O1) is selectively returned to +z)
;F2 and F3 decoupled
;using reduced presaturation power and offset jump to HN region
;in order to reduce sample heating

;A.G. Palmer III, J. Cavanagh, P.E. Wright & M. Rance, J. Magn.
;  Reson. 93, 151-170 (1991)
;L.E. Kay, P. Keifer & T. Saarinen, J. Am. Chem. Soc. 114,
;  10663-5 (1992)

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


;###########################
;#### IMPORTANT NOTES: ####
;#########################################################################
;## This interleaved sampling scheme creates a pseudo-3D spectrum
```

```
;## where F3(aq) and F2(td2) represent the known X,H-HSQC
;## The dummy -dimension F1(td1,l2) serves as an index for the
;## NOE or no-NOE spectra respectively.
;## Several pairs of NOE/no-NOE spectra can be acquired
;## according to the chosen value for TD1 (even numbers!!!)
;## PLEASE NOTE THAT, BEFORE PROCESSING,THE STATUS PARAMETER FOR TD1
;## MUST BE SET TO TD1 MANUALLY!!!
;## The efficient NOE build-up delay is d9, calculated from d30
;#############################################################################
;## The actual number of scans per experiment is not = NS, but L3=NS/TD1!
;## Thus,in the status window of the acquisition window,the scan counter
;## will only count up to l3 = NS/TD1 !
;## For same intensities, NS must be increased proportionally to TD1
;## For example: NS = 16*TD1 will accumulate 16 scans per experiment!
;#############################################################################

;########################################
;#### Important Processing Information ####
;#########################################################################################
;## BEFORE PROCESSING,THE STATUS PARAMETER FOR TD1 MUST BE SET TO TD1 MANU-
ALLY!
;## Only xfb in the F2/F3 plane meaningful!
;## td1 is index for NOE / no-NOE
;## MC2: echo-antiecho
;## Ph0(F1) = -90 degree
;#########################################################################################

;#### Calculation of Experiment Time ####
;## The 'expt' command yields too long experiment times
;## Calculate duration manually as: NS * 2TD * (d1+d30+AQ)

;################################################
;########### Setting Parameters: ################
;################################################
;FQ1LIST = contains 3 entries: BF(1H), O1(HN), O1 (H2O)
;SFO1= water on-resonant
;DS  = 2*n*NS
;td1/2 = number of pairs of NOE/no-NOE experiments
;in0 = 1/(2 * SW(X)) = DW(X)        ;d0 is here used for F2(td2)!!!
;nd0 = 2
;cpdprg2: garp
;GRDPROG: 2sineea(ma)
;use gradient ratios:
;           cnst21 : cnst22 : cnst23 : cnst24:
;15N        69           7           69      -7

;p1                                  ;90 H high power pulse on F1 (pl1)
;p3                                  ;90 X high power pulse on F2 (pl2)
;p10= ca.100u                        ;180 H low power pulse on F1 (pl11) for
                                     ;presaturation
;p16= ca. 1m                         ;gradient pulse length
;pcpd2                               ;90 low power pulse on F2 (pl12) for cpd2

;d1 = ca. 2*T1(HN)                   ;relaxation delay
;d4 =< 2.7m                          ;1/(4J(NH))
"d14=d4"                             ;change to 1/(6J(NH)) for NH2 groups
;d16=200u                           ;gradient recovery delay
;d29 = ca. 5m                       ;delay between hard pulses in presaturation loop
;d30 = ca. 3s                       ;presaturation time for NOE build-up
                                    ;the exact value will be d9!!!
```

```
;#### fixed parameters, do not change ####
"p2=p1*2"
"p4=p3*2"

"d0=3u"
"d11=50u"                              ;delay for phase and buffer incrementation
"d12=3u"                               ;delay for power level switching
"d13=3u"                               ;delay for phase presetting
"d6=(p4-p2)/2"                         ;180 pulse center delay
"d7=(p3-p1)/2"                         ;90 pulse center delay
"d24=d14+p1+d0"                        ;compensated d14
"d25=d13+p16+d16"                      ;compensated refocussing delay
"d15=d14-p16-d11-d0-p1"                ;compensated refocussing delay

"l0=(td2/2)"
"l1=d30/(d29+p10*0.67)+0.5"            ;loop counter for presaturation
"l3=ns/td1"                            ;reduced scan counter
"l8=1"                                 ;dummy scan flag
"l9=1"                                 ;flag for NOE/noNOE
"nbl=td1"                              ;number of buffers allocated

;#### Efficient NOE build/up time ####
"d9=l1*(d29+p10*0.67)"


 ze
 d11 st0
 d12 pl12:f2
1 zd
 d15 iu8                               ;increment dummy scan flag
2 d11*6
3 d1
 d11
4 d11*3
5 d12 pl11:f1
 5m fq1:f1                             ;set SFO1 to center of HN protons
 if "l9 == 2" goto 7                   ;if label l9=2 then no-NOE else NOE
 d13
6 d29
 p10*0.67 ph0                          ;proton presaturation loop
 lo to 6 times l1
 d12 pl2:f2 iu9
 goto 8
7 d9 ru9                               ;no-NOE part and reset NOE flag
 d12 pl2:f2
8 d13
 (p3 ph5):f2                           ;begin of pulse sequence - excitation pulse on N
 d24 fq1:f1                            ;refocussing of chemical shift evolution, jump to
                                       ;H2O
 (p4 ph4):f2
 d0 pl1:f1                             ;t1 chemical shift evolution
 (p2 ph8):f1
 d0
 d11 UNBLKGRAD
 p16:gp1                               ;select 15N-SQC
 d15
 (d7 p1 ph0) (p3 ph6):f2               ;INEPT transfer of x-component
 d14
 (d6 p2 ph0) (p4 ph0):f2
```

```
 d14
 (d7 p1 ph1) (p3 ph7):f2                ;INEPT transfer of y-component (Palmer et al)
 d4
 (d6 p2 ph0) (p4 ph0):f2
 d4
 d7
 (p1 ph2)                               ;read pulse
 d25
 (p2 ph0)
 d13
 p16:gp2*EA                             ;select 1H-SQC
 d16 pl12:f2
 goscnp ph31 cpd:f2                     ;acquire without loop and phase pointer
                                        ;incrementation
 d11 do:f2
 d11  BLKGRAD
 if "l8 < l5+2" goto 1                  ;conditional loop for DS dummy scans
 d11 st                                 ;increment buffer for next experiment
 lo to 2 times td1                      ;loop for next experiment same scan
 d11 ipp31                              ;increment phase program pointers
 d11 ipp4
 d11 ipp5
 d11 ipp6
 d11 ipp7
 d11 ipp8
 lo to 3 times l3                       ;loop for next scan (l3=ns/td1)
 d1 wr #0 if #0 zd                      ;write out buffers and move disk pointer by TD*nbl
 d11 ip6*2 igrad EA                     ;shift sign for sensitivity enhancement
 lo to 4 times 2                        ;loop once for anti-echo detection
 d11 id0
 d11 ip5*2
 d11 ip31*2
 lo to 5 times l0                       ;loop for (td2/2) complex time points in F2
exit


ph0= 0
ph1= 1
ph2= 2
ph4= 0 1 2 3
ph5= 1 1 3 3
ph6= 0 0 0 0 2 2 2 2
ph7= 1 1 1 1 3 3 3 3
ph8= 0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph31=0 2 2 0 2 0 0 2
```

Pseudo 3D MEXICO

```
;tdmexico.3d                    td 18/04/2000
;avance-version

;MEXICO sequence using
;2D H-1/X correlation via double inept transfer
;using water flip back technique for sensitivity enhancement
;of fast exchanging protons

;W.Jahnke, Dissertation, TU Muenchen, 1994

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

```
#include <Delay.incl>

;****Channel assignments****
;f1- 1H
;f2- 13Ca
;(f3- 13CO, this channel is not needed)
;f4- 15N
;use gradient program (GRDPROG) :          6sineeama
;use gradient ratios:
;-55 : -60 : 40 : -35 : -10 :-55 : -60 : 40 : -35 : 10          for 15N

;**** Parameters to be set ****
;p1                                        ;90 high power 1H (pl1,F1)
;p3 as short as possible                   ;90 high power Ca (pl2,F2)
;p7 >p4! (90deg 15N,pl4,F4)
;pcpd4                                     ;90 low power 15N (pl14.F4) for cpd dec.
;p16 >800u (gradient duration)
;d1 > 1s (relaxation delay)
;d16 >150u (field recovery)

;TD1 is the dummy dimension:
;TD1 = number of mixing times plus 1 (HSQC)
;    = number of entries in vdlist
;vdlist: contains TD1 entries:
;1. entry = any value (HSQC)
;2.+ entries = mixing times
;nd0=nd10=2 (echo/antiecho)
;nbl = td1! (number of experiments)

;**** calculated parameters ****
"p2=p1*2"                                 ;180 high power 1H
"p4=p3*2"                                 ;180 high power Ca
"p8=p7*2"                                 ;180 high power 15N
define delay cen26
define delay cen27
define delay cen28
define delay cen46
define delay cen47
"cen26=(p6-p2)/2"
"cen27=(p7-p2)/2"
"cen28=(p8-p2)/2"
"cen46=(p6-p4)/2"
"cen47=(p7-p4)/2"
"d0=3u"                                   ;incrementable t1 delay
"d2=5.5m"                                 ;1/2J(HN)
"d3=3.6m"                                 ;1/2J(CH) optimized for CaHa
"d11=10m"
"d20=p6+d0*2"
"d23=d2-d3-p3"
"d24=d4-p16"
"d25=d3-d23-p7"
"d26=d2-d25-p3"
"nbl=td1"
"l3=(td2/2)"
"l9=1"                                    ;flag

1 ze
  d11 st0
2 d11 do:f4
3 d1
```

```
  d11*2
4 d11*4
5 10u pl4:f4
  if "l9 == 1" goto 6                                   ;first ser is HSQC
;### Maximum Quantum Filter for 13C and 15N
  (p1 ph10):f1                                          ;H2O in -y
  d3
  (p3 ph11):f2
  d23
  (cen27 p2 ph0):f1 (cen47 p4 ph0):f2 (p7 ph11):f4
  d25
  (p3 ph12):f2
  d26                                                   ;H2O in y
  (p1 ph10):f1 (p7 ph12):f4                             ;H2O in z
  ;### H2O along z, all other dephased
6 50u UNBLKGRAD
  p16:gp1
  d16
  vd                                                    ;mixing time for H2O-HN magnetization
                                                        ;transfer
  ;### Start of ordinary FLIP-BACK 15N-HSQC{13C,13CO}
7 (p1 ph0):f1                                           ;H2O in -y
  d4
  (cen28 p2 ph1):f1 (p8 ph3):f4
  d4
  (p1 ph7):f1                                           ;H2O in -y
  p16:gp2
  d16
  (p7 ph3):f4                                           ;H2O in -y
  d0
  (cen26 p2 ph6):f1 (cen46 p4 ph0):f2 (p6:sp6 ph0):f3           ;H2O in -y
  d0
  p16:gp3
  d16
  (p8 ph4):f4
  p16:gp4
  d16
  d20
  (p7 ph0):f4
  p16:gp5
  d16                                                   ;H2O in -y
  (p1 ph0):f1                                           ;H2O in -z
  d4
  (cen28 p2 ph0):f1 (p8 ph0):f4                         ;H2O in +z
  p16:gp6*EA
  d24 pl14:f4
  4u  BLKGRAD
  go=2 ph31 cpd4:f4
  d11 do:f4 st                                          ;record echo parts
  d11 iu9                                               ;increment flag
  d11 ivd                                               ;increment mixing time
  lo to 3 times td1                                     ;loop for mixing times
  d1 wr #0 if #0 zd                                     ;write echo parts
  d11 ru9 st0                                           ;reset flag
  d11 igrad EA                                          ;switch to antiecho
  lo to 4 times 2                                       ;loop for echo/antiecho
  d11 ip31*2
  d11 ip3*2
  d11 id0
  lo to 5 times l3                                      ;loop for next t1 increment
```

```
exit

ph0=0
ph1=1
ph3=0 2                                   ;inverts, axial peak suppression in F1
ph4={{0}*2}^1                             ;inverts, phase purge in F1
ph6={{1}*2}^2                             ;suppresses MQC in F1
ph7={{1}*4}^2                             ;inverts
ph10={{0}*8}^2                            ;no effect
ph11={{0}*2}^2                            ;destructive filter
ph12={{0}*4}^2                            ;destructive filter
ph31=0 2 2 0 2 0 0 2
```

2D ^{13}C,H[^{12}C,^{14}N]-NOESY

```
;tdCH_noesy.wg                           td 08/02/2000
;avance-version

;2D (H)13C,H(12C,14N)-NOESY with C,H editing
;destructive 15N and 13C filter (double) in F2
;water suppression using 3-9-19 pulse sequence with gradients

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

;nd0 = 2 (STATES)
;reverse = FALSE in both dimensions
;water must be on-resonance (sfo1)
;sfo2 = center of 13C (ca. 45 ppm)
;sfo4 = center of 15N (ca. 120 ppm)
;use gradient program (GRDPROG) :        5sinema
;use gradient ratio:          cnst24 = cnst25, all other are variable z-spoils!
;cnst21:cnst22:cnst23:cnst24:cnst25
;  -33 :  70  :  15  :  20  :  20

;d1 : relaxation delay; 1-5 * T1
;d2 = 3.4 - 3.6m, suppression of H(13C)
;d3 = 1.3 - 1.8m (<= 1/4JCH)
;d4 = 5.0m, suppression of H(15N)
;d8 >50m (NOE mixing time)
;d16 = 200u (gradient recovery delay)
;d19 ca.60-120u (WATERGATE delay)
;NS: 8 * n
;p1 = 90deg 1H pulse (pl1,F1)
;p3 = 90deg 13C pulse (pl2=0dB,F2)
;p7 = 90deg 15N pulse(pl4,F4)

"p2=p1*2"
"p8=p7*2"
"p4=p3*2"
"l1=td1/2"
define delay wg
define delay cen2
define delay cen4
define delay cen8
"wg=p1*4.77+d19*10"
"cen2=(p4-p2)/2"
```

```
"cen4=(wg-p4)/2"
"cen8=(wg-p8)/2"
"d0=3u"
"d11=10m"
"d18=d8-p16-d16"
"d22=d2-p16-d16"
"d23=d4-d2"
"d24=d2*2-d4-p7-p16"



1 ze
2 d1
3 d11*2
4 (p1 ph0)                                  ;-H2Oy
  d3
  (cen2 p2 ph1) (p4 ph3):f2
  d3 UNBLKGRAD
  (p1 ph7)
  GRADIENT(cnst21)                          ;z-spoil
  d16
  (p3 ph3):f2
  d0
  (p2 ph6)
  d0
  (p3 ph0):f2
  GRADIENT(cnst22)                          ;z-spoil
  d16                                       ;-H2Oy
  (p1 ph1)                                  ;2CzHx
  d3
  (cen2 p2 ph0) (p4 ph0):f2                 ;H2Oy
  d3                                        ;Hy
  (p1 ph0)                                  ;H2Oz
;### NOE mixing ###
  d18
  GRADIENT(cnst23)                          ;z-spoil
  d16
;### destructive 13C,15N filter with WATERGATE ###
  (p1 ph10)
  d22 UNBLKGRAD
  GRADIENT(cnst24)
  d16
  (p3 ph10):f2
  (p1*0.231 ph4 d19*2 p1*0.692 ph4 d19*2 p1*1.462 ph4 d19*2 p1*1.462 ph5 d19*2 p1*0.692 ph5 d19*2
p1*0.231 ph5):f1 (cen4 p4 ph12):f2  (cen8 p8 ph12):f4
  d23
  (p7 ph11):f4
  GRADIENT(cnst25
  d24
  (p3 ph11):f2
  10u  BLKGRAD
  go=2 ph31
  d1 wr #0 if #0 ip3 zd
  lo to 3 times 2
  d11 id0
  d11 ip31*2
  lo to 4 times l1
exit

ph0 = 0
```

```
ph1 = 1
ph3 = 0 2                          ;constructive 13C filter and STATES phase
ph4 ={{0}*8}^1                     ;WATERGATE phase
ph5 ={{2}*8}^1                     ;WATERGATE phase
ph6 = 1 1 3 3                      ;MQC suppression in F1
ph7 = {{1}*4}^2                    ;water suppression in INEPT
ph10= 0 0 2 2                      ;inverts
ph11= 0 2                          ;destructive 15N+13C filter
ph12= {{0}*4}^1                    ;dito on 180deg pulses
ph31= 0 2 2 0 2 0 0 2 2 0 0 2 0 2 2 0
```

Pseudo 2D Diffusion experiment (on water)

```
;tddiff.2D                         td 13/03/2001
;avance-version

;with first order compensation of convection flux
;optional presaturation and WATERGATE suppression scheme
;Same program may be used to record diffusion of protein and water!

;#### ATTENTION #####
;Note a stupid BRUKER bug:
;under WIXNMR 2.6, the initialisation of the first gradient
;strength fails and produces an excessively strong FID which
;blows the receiver gain. Therefore, this program increments
;the gradient step list by one, and consequently, the last FID is
;actually the first FID. It must be deleted before processing
;by setting NUMPNTS in the 'edte1' setup to SI1 = TD1-1 !
;Once the bug has been fixed, delete annotated lines!

;Lit:
;A.Jerchow, N.Mueller in JMR 125 (1997), 372-375

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

;######## Important Notes #########
;## ALWAYS run diffusion experiments on the protein AND the water, i.e.
;## H2O: l0 = 0 AND pl11=120 (H2O, no presat and WATERGATE)
;## Protein: l0 = 1 AND pl11 ~55dB
;## in order to verify your results: D(H2O) is tabulated!
;## On changing from square to sine gradients (gp2),
;## multiply d8 by 2.5 to get equal I(gradient) curve
;## DUE TO THE BRUKER BUG EXPLAINED ABOVE, SET
;## TD1 = SI1+1, e.g. TD1 = 33 while SI1 = 32!

;### parameters to be set ###
;l0 = 0(no Watergate,pl11=120!),1(Watergate)
;gpx1 = 1.35*gpz1
;gpy1 = 1.35*gpz1
;gpnam2 = sine.100 or square.100
;gpz2 =< 60 (to reduce phase errors!)
;gpx2 = 0
;gpy2 = 0
;p1 = 90deg 1H pulse (pl1)
;pl1 >= 0dB (for p1=90deg)
;pl11 ~ 55dB presat power (120dB = no presat)
;p16 >= 800us (z-spoil gradient pulse)
```

```
;p17 ~ 2ms (diffusion-editing gradient)

;d1 ~1s relaxation and presat delay
;d8 = BD, ~100-200ms (protein), ~20ms (H2O)
;d9 >= 10m z-filter delay (LED)
;d17 ~ 2*p17 (recovery after diffusion gradients)
;d16 > 150u (recovery delay for z-spoil gradients)
;d19 > dw, WATERGATE suppression delay

;### calculated parameters ###

define delay LD
define delay BD
define delay wg
"p2=p1*2"
"d3=d8-p17*10/3-d17*3"
"d11=30m"
"d13=d3*0.5-p16-d16"
"d20=l0*10m+10m"
"wg=p1*4.77+d19*10"
lgrad step = td1

;## Calculating Diffusion Coefficients in XWINNMR T1/T2 Mode ###
;  Setup with 'edt1':
"LD=p17*2"
"BD=d8"
;function type = vargrad: I(grad)=I0*exp[-1e-5*D*(BD-LD/3)*(2PI*g*LD*grad)^2]
;g = gamma(1H) = 4258*gscal*gshape (in G/cm)
;where    gshape = 1 for square grads and 2/PI for sine grads
;            gscal = calibrated gradient strength factor (0.67 for TXI on PAT)
;LISTTYPE = auto
;list increment = gp2z/SI1

1 ze
  LD UNBLKGRAD
  d20 igrad step                          ;delete this lines once BRUKER bug is fixed!
  p17:gp2*step                            ;delete this lines once BRUKER bug is fixed!
  BD
2 d11*2  BLKGRAD
3 10u pl11:f1
  d1 cw:f1                                ;H2O presaturation
  10u do:f1
  50u pl1:f1 UNBLKGRAD
;### First dephasing DPFGE ###
  p1 ph1
  p17:gp2*step
  d17
  p2 ph2
  p17:gp2*-1.0*step
  d17
  p1 ph3
;### First half mixing time ###
  d13
  p16:gp1*-0.7                            ;-0.7
  d16
;### Concatenated first rephasing and second dephasing DPFGE ###
  p1 ph4
  p17:gp2*step
  d17
  p17:gp2*step
```

```
  d17
  p2 ph5
  p17:gp2*-1.0*step
  d17
  p17:gp2*-1.0*step
  d17
  p1 ph6
;### Second half mixing time ###
  d13
  p16:gp1*-0.3                              ;*-0.3
  d16
;### Second rephasing DPFGE ###
  p1 ph7
  p17:gp2*step
  d17
  p2 ph8
  p17:gp2*-1.0*step
  d17
  p1 ph11
;### LED (= z-filter) ###
  p16:gp1*-0.5                              ;*-0.5
  d16
  d9
;### Read, alternatively with WATERGATE ###
  p1 ph10
  p16:gp1
  d16
 if(l0==0)
{
 (p2 ph15)
}
else
{
 (p1*0.231 ph13 d19*2 p1*0.692 ph13 d19*2 p1*1.462 ph13 d19*2 p1*1.462 ph14 d19*2 p1*0.692 ph14
d19*2 p1*0.231 ph14):f1
}
  p16:gp1
  d16
  go=2 ph31
  d11  BLKGRAD wr #0 if #0 zd
  d11 igrad step
  lo to 3 times td1
 exit

ph0=0
ph1=0 0 1 1 2 2 3 3
ph2=1 2 2 3 3 0 0 1
ph3={2 0 3 1 0 2 1 3}^2
ph4=1 1 2 2 3 3 0 0
ph5=0 1 1 2 2 3 3 0
ph6={{3 1 0 2 1 3 2 0}*2}^2
ph7=0 2
ph8=1 1 2 2 3 3 0 0
ph10=0 2
ph11=0 0 2 2
ph13=0 0 1 1 2 2 3 3
ph14=2 2 3 3 0 0 1 1
ph15=0 0 1 1
ph31={{{0 0 2 2 }*2}^2}^2
```

Pseudo 2D Diffusion experiment (on proteins)

```
;tddiff.2Dwg                                    td 13/03/99
;avance-version

;with first order compensation of convection flux
;water suppression using WATERGATE

;Lit:
;A.Jerchow, N.Mueller in JMR 125 (1997), 372-375


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

;######## Important Notes #########
;grdprog = tddiff2Dwg (square gradients) or tddiff2Dswg (sinus gradients)
;only z-spoil gradients matter, e.g.:
;cnst22 = 19, cnst23 = 10 cnst24 = 12 proposed
;set L20 = DS!!
;start the experiment not with zg, but with dosy!

;###### Processing Info #####
;only do xf2, no zero-filling!
;1 SI = 1 TD

;###### Analysis (BRUKER T1/T2 menu) #####
;listtyp = auto
;fit function = vargrad
;Little Delta = LD, Big Delta = BD
;Set proper scaling with gamma:
;gamma = 4258*s*z
;s = 1 for square gradients, 2/PI for sinus gradients
;z = gradient scaling factor as calibrated, i.e. 0.67 for TXI/DMX601

;**** parameters to be set *****
;p0                                     ;90degree proton pulse (pl18) for fine tuning of WATERGATE
;p1                                     ;90degree proton pulse (pl1)
;p16 ca. 1ms                            ;gradient pulse length for z-spoils
;p17 ca. 2-3ms                          ;gradient pulse length for encoding rectangular gradients
(cnst21)
;p28                                    ;90degree proton pulse (pl18) for WATERGATE

;d1                                     ;relaxation and presaturation (pl11)
;d3 ca. 20-50ms                         ;mixing time for diffusion
;d17 > p17                              ;gradient recovery delay after square pulses
"d18=d17"                               ;gradient recovery delay after square pulses
;d16 > 100us                            ;sine-shaped gradient recovery delay
;d19 >= dw                              ;WATERGATE suppression delay


;**** calculated parameters ****
define delay LD
define delay BD
"p2=p1*2"
"d13=d3*0.5-p16-d16"
"d23=(d3+d17-d18)/2-p16-d16"
"ds=l20"
```

```
;#####Fit Function as provided by XWINNMR (type=vargrad)#########

;####  I(cnst21)=I0*exp[-1e-5*D*(BD-LD/3)*(2PI*gamma*LD*cnst21)^2]
"LD=p17*2"
"BD=d3+p17*10/3+d17*5/2+d18/2"


1 ze
  LD
  BD
2 d1
3 50u UNBLKGRAD
  p1 ph1                                ;start of first PFGE
   p17:ngrad:c34
  d17
  p2 ph2
   p17:ngrad:c34
  d17
  p1 ph3                                ;end of first PFGE
  d13                                   ;start of first diffusion mixing
   p16:ngrad:c34
  d16
  p1 ph4                                ;start of rephasing PFGE
   p17:ngrad:c34
  d18
   p17:ngrad:c34
  d18
  p2 ph5
   p17:ngrad:c34
  d18
   p17:ngrad:c34
  d18
  p1 ph6                                ;end of rephasing PFGE
  d23                                   ;start of second diffusion mixing
   p16:ngrad:c34
  d16
  p1 ph7                                ;start of second PFGE
   p17:ngrad:c34
  d17
  p2 ph8
   p17:ngrad:c34
  d17

  4u
   p16:ngrad:c34                        ;start of WATERGATE and end of second PFGE
  d16 pl18:f1
  p28*0.231 ph13
  d19*2
  p28*0.692 ph13
  d19*2
  p28*1.462 ph13
  d19*2
  p28*1.462 ph14
  d19*2
  p28*0.692 ph14
  d19*2
  p0*0.231 ph14
   p16:ngrad:c34
  d16 pl1:f1
```

```
 4u  BLKGRAD

 go=2 ph31
 d1 wr #0 if #0 zd
 lo to 3 times td1
exit

ph1=0 0 1 1 2 2 3 3
ph2=1 2 2 3 3 0 0 1
ph3=0
ph4=2 2 3 3
ph5=3 3 0 0 3 3 0 0 0 0 1 1 0 0 1 1
ph6=2
ph7=0
ph8={0}*16 {2}*16
ph13= {0}*16 {1}*16 {2}*16 {3}*16
ph14= {2}*16 {3}*16 {0}*16 {1}*16
ph31=0 2 0 2 2 0 2 0 2 0 2 0 0 2 0 2
   2 0 2 0 0 2 0 2 0 2 0 2 2 0 2 0
```

7 References

[1] L. Stryer, *Biochemistry*, 3rd ed., W. H. Freeman and Company, New York **1988**.

[2] T. E. Creighton, *Proteins*, W. H. Freeman and Company, New York **1993**.

[3] G. Wider, K. Wüthrich, *Curr. Op. Struct. Biol* **1999**, *9*, 594-601.

[4] X. Shan, K. H. Gardner, D. R. Muhandiram, N. S. Rao, C. H. Arrowsmith, L. E. Kay, *J. Am. Chem. Soc* **1996**, *118*, 6570-6579.

[5] L. E. Kay, K. H. Gardner, *Curr. Op. Struct. Biol* **1997**, *7*, 722-731.

[6] G. M. Clore, A. M. Gronenborn, *Curr. Op. Struct. Biol* **1998**, *2*, 564-570.

[7] S. B. Shuker, P. J. Hadjuk, R. P. Meadows, S. W. Fesik, *Science* **1996**, *274*, 1531-1534.

[8] P. J. Hajduk, G. Sheppard, D. G. Nettesheim, E. T. Olejniczak, S. B. Shuker, R. P. Meadows, D. H. Steinman, G. M. Carrera, P. A. Marcotte, J. Severin, K. Walter, H. Smith, E. Gubbins, R. Simmer, T. F. Holzman, D. W. Morgan, S. K. Davidsen, J. B. Summers, S. W. Fesik, *J. Am. Chem. Soc* **1997**, *119*, 5818-5827.

[9] P. J. Hadjuk, T. Gerfin, J. M. Boehlen, M. Häberli, D. Marek, S. W. Fesik, *J. Med. Chem* **1999**, *42*, 2315-2317.

[10] S. Eberhardt, G. Richter, W. Gimbel, T. Werner, A. Bacher, *Eur. J. Biochem.* **1996**, *242*, 712-719.

[11] F. Bloch, W. W. Hansen, *Phys. Rev.* **1946**, *69*, 127.

[12] E. M. Purcell, H. C. Torrey, R. V. Pound, *Phys. Rev.* **1946**, *69*, 37-38.

[13] R. R. Ernst, W. A. Anderson, *Rev. Sci. Instr.* **1966**, *37*, 93-102.

[14] R. R. Ernst, *Angew. Chem. Int. Ed. Engl* **1992**, *31*, 805-823.

[15] W. P. Aue, E. Bartholdi, R. R. Ernst, *J. Chem. Phys* **1976**, *64*, 2229-2246.

[16] R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford **1987**.

[17] L. Braunschweiler, R. R. Ernst, *J. Magn. Reson* **1983**, *53*, 521-528.

[18] J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys* **1979**, *71*, 4546-4553.

[19] A. Kumar, R. R. Ernst, K. Wüthrich, *Biochem. Biophys. Comm* **1980**, *95*, 1-6.

[20] V. F. Bystrov, *Prog. Nucl. Magn. Reson. Spect* **1976**, *10*, 41-81.

[21] K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York **1986**.

[22] A. A. Bothner-By, R. L. Stephens, J. Lee, C. D. Warren, R. W. Jeanloz, *J. Am. Chem. Soc* **1984**, *106*, 811-813.

[23] A. Bax, D. G. Davis, *J. Magn. Reson* **1985**, *63*, 207-213.

[24] G. W. Vuister, R. Boelens, *J. Magn. Reson* **1987**, *73*, 328-333.

[25] G. W. Vuister, R. Boelens, R. Kaptein, *J. Magn. Reson* **1988**, *80*, 176-185.

[26] C. Griesinger, O. W. Sorensen, R. R. Ernst, *J. Magn. Reson* **1987**, *73*, 574-579.

[27] C. Griesinger, O. W. Sorensen, R. R. Ernst, *J. Am. Chem. Soc* **1987**, *109*, 7227-7228.

[28] H. Oschkinat, C. Griesinger, P. J. Kraulis, O. W. Sorensen, R. R. Ernst, A. M. Gronenborn, G. M. Clore, *Nature* **1988**, *332*, 374-376.

[29] L. E. Kay, G. M. Clore, A. Bax, A. M. Gronenborn, *Science* **1990**, *249*, 411-414.

[30] D. Marion, L. E. Kay, S. W. Sparks, D. A. Torchia, A. Bax, *J. Am. Chem. Soc* **1989**, *111*, 1515-1517.

[31] S. W. Fesik, E. R. P. Zuiderweg, *J. Magn. Reson* **1988**, *78*, 588-593.

[32] L. P. McIntosh, F. W. Dahlquist, *Quart. Rev. Biophys* **1990**, *23*, 1-38.

[33] S. L. Schreiber, G. L. Verdine, *Tetrahedron* **1991**, *47*, 2543-2562.

[34] D. Marion, P. C. Driscoll, L. E. Kay, P. T. Wingfield, A. Bax, A. M. Gronenborn, G. M. Clore, *Biochemistry* **1989**, *29*, 150-156.

[35] G. A. Morris, R. Freeman, *J. Am. Chem. Soc* **1979**, *101*, 760-762.

[36] D. H. Live, D. Davis, W. C. Agosta, D. Cowburn, *J. Am. Chem. Soc* **1984**, *106*, 6104-6105.

[37] I. Solomon, *Phys. Rev.* **1955**, *99*, 559.

[38] D. Neuhaus, M. Williamson, *The Nuclear Overhauser Effect in Sturctural and Conformational Analysis*,VCH Weinheim, New York **1989**.

[39] R. Brüschweiler, R. R. Ernst, *J. Chem. Phys* **1991**, *96*, 1758-1766.

[40] R. Riek, G. Wider, K. Pervushin, K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4918-4923.

[41] H. Kessler, C. Griesinger, J. Zarbock, H. R. Loosli, *J. Magn. Reson* **1984**, *57*, 331-336.

[42] H. Kessler, C. Griesinger, G. Zimmermann, *Magn. reson. chem* **1987**, *25*, 579-583.

[43] G. W. Vuister, A. Bax, *J. Magn. Reson* **1992**, *98*, 428-435.

[44] T. M. Logan, E. T. Olejniczak, R. X Xu, S. W. Fesik, *J. Biomol. NMR* **1993**, *3*, 225-231.

[45] S. Grzesiek, J. Anglister, A. Bax, *J. Magn. Reson* **1993**, *B 101*, 114-119.

[46] D. M. LeMaster, *Annu. Rev. Biophys. Chem* **1990**, *19*, 243-266.

[47] P. E. Coughlin, F. E. Anderson, E. J. Oliver, J. M. Brown, S. W. Homans, S. Pollak, J. W. Lustbader, *J. Am. Chem. Soc* **1999**, *121*, 11871-11874.

[48] N. K. Gotho, L. E. Kay, *Curr. Op. Struct. Biol* **2000**, *10*, 585-592.

[49] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12366-12371.

[50] M. Salzmann, G. Wider, K. Pervushin, H. Senn, K. Wüthrich, *J. Am. Chem. Soc* **1999**, *121*, 844-848.

[51] V. Dötsch, G. Wagner, *Curr. Op. Struct. Biol* **1998**, *8*, 619-623.

[52] M. Tessari, F. A. A. Mulder, R. Boelens, G. W. Vuister, *J. Magn. Reson* **1997**, *127*, 128-133.

[53] M. Tessari, H. Vis, R. Boelens, R. Kaptein, G. W. Vuister, *J. Am. Chem. Soc* **1997**, *119*, 8985-8990.

[54] J. C. Hus, D. Marion, M. Blackledge, *J. Am. Chem. Soc* **2001**, *123*, 1541-1542.

[55] M. Ikura, L. E. Kay, A. Bax, *Biochemistry* **1990**, *29*, 4659- 4667.

[56] L. E. Kay, M. Ikura, R. Tschudin, A. Bax, *J. Magn. Reson* **1990**, *89*, 496-514.

[57] A. Bax, M. Ikura, *J. Biomol. NMR* **1991**, *1*, 99-104.

[58] G. W. Vuister, A. Bax, *J. Am. Chem. Soc* **1993**, *115*, 7772-7777.

[59] M. Wittekind, L. Mueller, *J. Magn. Reson* **1993**, *101(B)*, 201-205.

[60] S. Grzesiek, A. Bax, *J. Am. Chem. Soc* **1992**, *114*, 6291-6293.

[61] S. Seip, J. Balbach, H. Kessler, *J. Magn. Reson* **1992**, *100*, 406-410.

[62] R. T. Clubb, V. Thanabal, G. Wagner, *J. Biomol. NMR* **1992**, *2*, 203-210.

[63] S. J. Archer, M. Ikura, D. A. Torchia, A. Bax, *J. Magn. Reson* **1991**, *95*, 636-641.

[64] M. Nilges, M. J. Macias, S. O'Donoghue, H. Oschkinat, *J. Mol. Biol.* **1997**, *269*, 408-422.

[65] D. Croft, J. Kemmink, K. P. Neidig, H. Oschkinat, *J. Biomol. NMR* **1997**, *10*, 207-219.

[66] M. Leutner, R. M. Gschwind, J. Liermann, C. Schwarz, G. Gemmecker, H. Kessler, *J. Biomol. NMR* **1998**, *11*, 31-43.

[67] C. Mumenthaler, P. Güntert, W. Braun, K. Wüthrich, *J. Biomol. NMR* **1997**, *10*, 351-362.

[68] E. T. Olejniczak, S. W. Fesik, *J. Am. Chem. Soc* **1994**, *116*, 2215-2216.

[69] M. Sattler, J. Schleucher, C. Griesinger, *Prog. Nucl. Magn. Reson. Spect* **1999**, *34*, 93-158.

[70] L. E. Kay, M. Ikura, A. Bax, *J. Am. Chem. Soc* **1990**, *112*, 888-889.

[71] G. T. Montelione, B. A. Lyons, S. D. Emerson, M. Tashiro, *J. Am. Chem. Soc* **1992**, *114*, 10974-10975.

[72] T. M. Logan, E. T. Olejniczak, R. X. Xu, S. W. Fesik, *FEBS Lett* **1992**, *314*, 413-418.

[73] M. Ikura, L. E. Kay, A. Bax, *J. Biomol. NMR* **1991**, *1*, 299-304.

[74] S. J. Archer, M. Ikura, D. A. Torchia, A. Bax, *J. Magn. Reson* **1991**, *95*, 636-641.

[75] G. M. Clore, A. M. Gronenborn, *Crit. Rev. Biochem. Mol. Biol.* **1989**, *24*, 479-557.

[76] D. Neri, T. Szyperski, G. Otting, H. Senn, K. Wüthrich, *Biochemistry* **1989**, *28*, 7510-7516.

[77] T. Szyperski, D. Neri, B. Leiting, G. Otting, K. Wüthrich, *J. Biomol. NMR* **1992**, *2*, 323-334.

[78] R. K. Hill, S. Yan, S. M. Arfin, *J. Am. Chem. Soc* **1973**, *95*, 7857-7859.

[79] S. R. Sylvester, C. M. Stevens, *Biochemistry* **1979**, *18*, 4529-4531.

[80] M. Kainosho, k. Ajisaka, *J. Am. Chem. Soc* **1975**, *97*, 5630-5631.

[81] R. W. Curley Jr., M. J. Panigot, A. P. Hansen, S. W. Fesik, *J. Biomol. NMR* **1994**, *4*, 335-340.

[82] H. Vis, R. Boelens, M. Mariani, R. Stroop, C. E. Vorgias, K. S. Wilson, R. Kaptein, *Biochemistry* **1994**, *33*, 14858-14870.

[83] L. P. McIntosh, E. Brun, L. E. Kay, *J. Biomol. NMR* **1997**, *9*, 306-312.

[84] F. Löhr, H. Rüterjans, *J. Magn. Reson* **1997**, *124*, 255-258.

[85] C. Griesinger, O. W. Sorensen, R. R. Ernst, *J. Am. Chem. Soc* **1985**, *107*, 6394-6396.

[86] J. R. Tolman, J. H. Prestegard, *J. Magn. Reson* **1996**, *B112*, 269-274.

[87] K. Wüthrich, M. Billeter, W. Braun, *J. Mol. Biol.* **1983**, *169*, 949-961.

[88] D. S. Wishart, B. D. Sykes, F. M. Richards, *Biochemistry* **1992**, *31*, 1647-1651.

[89] D. S. Wishart, B. D. Sykes, *J. Biomol. NMR* **1994**, *4*, 171-180.

[90] M. D. Reily, V. Thanabal, D. O. Omecinsky, *J. Am. Chem. Soc* **1992**, *114*, 6251-6252.

[91] D. S. Wishart, B. D. Sykes, F. M. Richards, *J. Mol. Biol.* **1991**, *222*, 311-333.

[92] G. P. Gippert, P. F. Yip, P. E. Wright, D. A. Case, *Biochem. Pharmacol* **1990**, *40*, 15-22.

[93] G. Merutka, H. J. Dyson, P. E. Wright, *J. Biomol. NMR* **1995**, *5*, 14-24.

[94] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biomol. NMR* **1995**, *5*, 67-81.

[95] M. Karplus, *J. Phys. Chem* **1959**, *30*, 11-15.

[96] M. Karplus, *J. Am. Chem. Soc* **1963**, *85*, 2870-2871.

[97] S. Spera, A. Bax, *J. Am. Chem. Soc* **1991**, *113*, 5490-5492.

[98] F. Löhr, H. Rüterjans, *J. Biomol. NMR* **1995**, *5*, 25-36.

[99] B. Whitehead, C. J. Craven, J. P. Waltho, *Methods. Mol. Biol.* **1997**, *60*, 29-52.

[100] D. R. Muhandiram, N. A. Farrow, G. Y. Xu, S. H. Smallcombe, L. E. Kay, *J. Magn. Reson. Ser. B* **1993**, *102*, 317-321.

[101] O. Zhang, J. D. Forman-Kay, *Biochemistry* **1997**, *36*, 3959-3970.

[102] T. Diercks, M. Coles, H. Kessler, *J. Biomol. NMR* **1999**, *15*, 177-180.

[103] S. Grzesiek, P. Wingfield, S. Stahl, J. D. Kaufman, A. Bax, *J. Am. Chem. Soc* **1995**, *117*, 9594-9595.

[104] J. N. Emsley, J. C. Lindon, *Pergamon Press, Oxford* **1975**.

[105] N. Tjandra, A. Bax, *Science* **1997**, *278*, 1111-1114.

[106] S. Cavagnero, H. J. Dyson, P. E. Wright, *J. Biomol. NMR* **1999**, *13*, 387-391.

[107] M. R. Hansen, L. Mueller, A. Pardi, *Nat. Struct. Biol* **1998**, *5*, 1065-1074.

[108] G. M. Clore, M. R. Starich, A. M. Gronenborn, *J. Am. Chem. Soc* **1998**, *120*, 10571-10572.

[109] B. W. Koenig, J. Hu, M. Ottiger, S. Bose, R. W. Hendler, A. Bax, *J. Am. Chem. Soc* **1999**, *121*, 1385-1386.

[110] J. Sass, F. Cordier, A. Hoffmann, M. Rogowski, A. Cousin, J. G. Omichinski, H. Löwen, S. Grzesiek, *J. Am. Chem. Soc* **1999**, *121*, 2047-2055.

[111] R. S. Prosser, J. A. Losonczi, I. V. Shiyanovskaya, *J. Am. Chem. Soc* **1998**, *120*, 11010-11011.

[112] N. Tjandra, S. Grzesiek, A. Bax, *J. Am. Chem. Soc* **1996**, *118*, 6264-6272.

[113] N. Tjandra, A. Szabo, A. Bax, *J. Am. Chem. Soc* **1996**, *118*, 6986-6991.

[114] D. A. Torchia, S. W. Sparks, A. Bax, *J. Am. Chem. Soc* **1988**, *110*, 2320-2321.

[115] M. Coles, T. Diercks, J. Liermann, A. Gröger, B. Rockel, W. Baumeister, K. K. Koretke, A. Lupas, J. Peters, H. Kessler, *Current Biology* **1999**, *9*, 1158-1168.

[116] R. T. Clubb, J. G. Omichinski, G. M. Clore, A. M. Gronenborn, *FEBS Lett* **1994**, *338*, 93.

[117] M. Wittekind, C. Mapelli, B. T. Farmer II, K.-L. Suen, V. Goldfarb, J. Tsao, T. Lavoie, M. Barbacid, C. A. Meyers, L. Mueller, *Biochemistry* **1994**, *33*, 13531-13539.

[118] P. Neri, R. Meadows, G. Gemmecker, E. Olejniczak, D. Nettesheim, T. Logan, R. Simmer, R. Helfrich, T. Holzman, J. Severin, S. Fesik, *FEBS Lett* **1991**, *294*, 81.

[119] W. J. Metzler, A. J. Bell, E. Ernst, T. B. Lavoie, L. Mueller, *J. Biol. Chem* **1994**, *269*, 4620.

[120] K. Newkirk, W. Feng, W. Jiang, R. Tejero, S. D. Emerson, M. Inouye, G. T. Montelione, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5114.

[121] H. Günther, *NMR Spectroscopy* **1980**, *John Wiley and Sons*, New York.

[122] S. Linse, O. Teleman, T. Drakenberg, *Biochemistry* **1990**, *29*, 5925-5934.

[123] G. Gemmecker, W. Jahnke, H. Kessler, *J. Am. Chem. Soc.* **1993**, *115*, 11620-11621.

[124] A. G. Palmer III, *Curr. Op. Struct. Biol* **1997**, *7*, 732-737.

[125] N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, L. E. Kay, *Biochemistry* **1994**, *33*, 5984-6003.

[126] B. J. Stockman, *Prog. Nucl. Magn. Reson. Spect* **1998**, *33*, 109-151.

[127] A. L. Breeze, *Prog. Nucl. Magn. Reson. Spect* **2000**, *36*, 323-372.

[128] O. W. Sorensen, G. W. Eich, M. H. Levitt, G. Bodenhausen, R. R. Ernst, *Prog. Nucl. Magn. Reson. Spect* **1983**, *16*, 163.

[129] E. Wörgötter, G. Wagner, K. Wüthrich, *J. Am. Chem. Soc* **1986**, *108*, 6162-6167.

[130] G. Otting, H. Senn, G. Wagner, K. Wüthrich, *J. Magn. Reson* **1986**, *70*, 500-505.

[131] H. Kogler, O. W. Sorensen, G. Bodenhausen, R. R. Ernst, *J. Magn. Reson* **1983**, *55*, 157-163.

[132] G. Gemmecker, E. T. Olejniczak, S. W. Fesik, *J. Magn. Reson* **1992**, *96*, 199-204.

[133] K. Ogura, H. Terasawa, F. Inagaki, *J. Biomol. NMR* **1996**, *8*, 492-498.

[134] R. H. A. Folmer, C. W. Hilbers, R. N. H. Konings, K. Hallenga, *J. Biomol. NMR* **1995**, *5*, 427-432.

[135] H. Geen, R. Freeman, *J. Magn. Reson* **1991**, *93*, 93-141.

[136] C. Zwahlen, P. Legault, S. J. F. Vincent, J. Greenblatt, R. Konrat, L. E. Kay, *J. Am. Chem. Soc* **1997**, *119*, 6711-6721.

[137] S. C. Li, C. Zwahlen, S. J. F. Vincent, C. J. McGlade, L. E. Kay, T. Pawson, J. D. Forman-Kay, *Nat. Struct. Biol* **1998**, *5*, 1075.

[138] M. Nilges, *Proteins: Struct. Funct. Gen.* **1993**, *17*, 297-309.

[139] http://www.caducee.net/Fiches-techniques/vitamineB-2.asp

[140] http://medecinesnaturelles.com/pages/sante/vitamine/B2.html

[141] R. Kuhn, H. Rudy, T. Wagner-Jauregg, *Ber. dtsch.chem.Ges* **1933**, *66*, 1950-1956.

[142] R. Kuhn, K. Reinemund, H. Kaltschmitt, R. Ströbele, H. Trischmann, *Naturwissenschaften* **1935**, *23*, 260.

[143] W. H. Peterson, M. S. Peterson, *Bact. rev* **1945**, *9*, 49.

[144] D. J. O'Kane, B. Woodward, J. Lee, D. C. Prasher, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 1100-1104.

[145] D. J. O'Kane, D. C. Prasher, *Mol. Microbiol* **1992**, *6*, 443-449.

[146] A. Bacher, S. Eberhardt, G. Richter, in *Escherichia and Salmonella* (Ed.: F. C. Neidhardt), Amer. Soc. for Microbiology, Washington, DC, **1996**, pp. 657-664.

[147] H. Wacker, R. A. Harvey, C. H. Winestock, G. W. E. Plaut, *J. Biol. Chem.* **1964**, *239*, 3493-3497.

[148] R. Beach, G. W. E. Plaut, *J. Am. Chem. Soc.* **1970**, *92*, 2913-2916.

[149] T. Paterson, H. C. S. Wood, *J. Chem. Soc. Chem. Commun.* **1969**, 290-291.

[150] H. Sedlmaier, F. Müller, P. J. Keller, A. Bacher, *Zeitschr. Naturforsch.* **1987**, *42c*, 425-429.

[151] A. Bacher, H. C. Ludwig, H. Schnepple, Y. Ben-Shaul, *J. Mol. Biol.* **1986**, *187*, 75-86.

[152] R. Ladenstein, B. Meyer, R. Huber, H. Labischinski, K. Bartels, H. D. Bartunik, L. Bachmann, H. C. Ludwig, A. Bacher, *J. Mol. Biol.* **1986**, *187*, 87-100.

[153] R. Ladenstein, K. Ritsert, R. Huber, G. Richter, A. Bacher, *Eur. J. Biochem.* **1994**, *223*, 1007-1017.

[154] R. Ladenstein, M. Schneider, R. Huber, H. D. Bartunik, K. Wilson, K. Schott, A. Bacher, *J. Mol. Biol.* **1988**, *203*, 1045-1070.

[155] K. Ritsert, R. Huber, D. Turk, R. Ladenstein, K. Schmidt-Bäse, A. Bacher, *J. Mol. Biol.* **1995**, *253*, 151-167.

[156] A. Bacher, R. Baur, U. Eggers, H. D. Harders, M. K. Otto, H. Schnepple, *J. Biol. Chem* **1980**, *255*, 632-637.

[157] F. Doignon, N. Biteau, M. Crouzet, M. Aigle, *Yeast* **1993**, *9*, 189-199.

[158] S. Eberhardt, G. Richter, H. Ritz, J. Brandt, A. Bacher, *in Yagi, K.* **1994**, 63-66.

[159] K. Schott, J. Kellermann, F. Lottspeich, A. Bacher, *J. Biol. Chem.* **1990**, *265*, 4204-4209.

[160] W. Meining, G. Tibbelin, R. Ladenstein, S. Eberhardt, M. Fischer, A. Bacher, *J. Struct. Biol.* **1998**, *121*, 53-60.

[161] S. Eberhardt, N. Zingler, K. Kemter, G. Richter, M. Cushman, A. Bacher, *Eur. J. Biochem.* **2001**, *268*, 4315-4323.

[162] A. T. Brunger, *X-PLOR, Version 3.1. A system for X-ray crystallography and NMR.*, Yale University Press, New Haven, Connecticut, **1992**.

[163] P. Stilbs, *Prog. Nucl. Magn. Reson. Spect* **1987**, *19*, 1-45.

[164] H. Carr, E. M. Purcell, *Phys. Rev.* **1954**, *94*, 630.

[165] S. Meiboom, D. Gill, *Rev. Sci. Instr.* **1958**, *29*, 688.

[166] E. O. Stejskal, J. E. Tanner, *J. Chem. Phys* **1965**, *42*, 288-292.

[167] J. E. Tanner, *J. Chem. Phys* **1970**, *52*, 2523-2526.

[168] S. J. Gibbs, C. S. Johnson, Jr, *J. Magn. Reson* **1991**, *93*, 395-402.

[169] M. Coles, T. Diercks, B. Muehlenweg, S. Bartsch, V. Zölzer, H. Tschesche, H. Kessler, *J. Mol. Biol.* **1999**, *289*, 139-157.

[170] M. Coles, T. Diercks, J. Liermann, A. Gröger, A. Lupas, J. Peters, C. Koretke, B. Rockel, W. Baumeister, H. Kessler, *Curr. Biol* **1999**, *9*, 1158-1168.

[171] J. Kuszewski, A. M. Gronenborn, G. M. Clore, *Prot. Sci.* **1996**, *5*, 1067-1080.

[172] J. Kuszewski, A. M. Gronenborn, G. M. Clore, *J. Magn. Reson.* **1997**, *125*, 171-177.

[173] L. Holm, C. Sander, *J. Mol. Biol.* **1993**, *233*, 123-138.

[174] M. Ingelman, V. Bianchi, H. Eklund, *J. Mol. Biol.* **1997**, *268*, 147-157.

[175] M. Wang, D. L. Roberts, R. Paschke, T. M. Shea, B. S. Masters, J. J. Kim, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 8411-8416.

[176] U. Ermler, R. A. Siddiqui, R. Cramm, B. Friedrich, *Embo. J* **1995**, *14*, 6067.

[177] C. C. Correll, C. J. Batie, D. P. Ballou, M. L. Ludwig, *Science* **1992**, *258*, 1604-1610.

[178] S. Koide, W. Jahnke, P. E. Wright, *J. Biomol. NMR* **1995**, *6*, 306-312.

[179] M. Piotto, V. Saudek, V. Sklenar, *J. Biol. NMR* **1992**, *2*, 661-665.

[180] R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *Journal of Applied Crystallography* **1993**, *26*, 283-289.

[181] P. J. Kraulis, *J. Appl. Crystallogr.* **1991**, *24*, 946-950.

[182] E. A. Merritt, M. E. P. Murphy, *Acta Crystallogr.* **1994**, *D50*, 869-873.

[183] G. W. E. Plaut, R. Beach, in *Chemistry and biology of the pteridines.* (Ed.: W. Pfleiderer), Walter de Gruyter, Berlin, **1975**, pp. 101-123.

[184] K. Otto, A. Bacher, *Biochemistry* **1981**, *115*, 511-517.

[185] V. Truffault, M. Coles, T. Diercks, K. Abelmann, S. Eberhardt, H. Lüttgen, A. Bacher, H. Kessler, *J. Mol. Biol.* **2001**, *309*, 949-960.

[186] D. I. Liao, Z. Wawrzak, J. C. Calabrese, P. V. Viitanen, D. B. Jordan, *Structure* **2001**, *9*, 399-408.

[187] G. W. E. Plaut, R. L. Beach, in *Flavins and flavoproteins*. (Ed.: T. P. Singer), Elsevier Scientific Publishing, Amsterdam, **1976**, pp. 737-746.

[188] R. Beach, G. W. E. Plaut, *Biochemistry* **1970**, *9*, 760-770.

[189] G. W. E. Plaut, R. Beach, T. Aogaichi, *Biochemstry* **1970**, *9*, 771-785.

[190] D. H. Brown, P. J. Keller, H. G. Floss, H. Sedlmaier, A. Bacher, *J. Org. Chem.* **1986**, *51*, 2461-2467.

[191] B. Illarionov, W. Eisenreich, A. Bacher, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 7224-7229.

[192] W. Kenji, M. Takashi, I. Akitami, O. Hideaki, H. Mamoru, *J. Biol. Chem* **2000**, *275*, 38393-38401.

[193] K. Auclair, A. Sutherland, J. Kennedy, D. J. Witter, J. P. Van den Heever, C. Richard Hutchinson, J. C. Vederas, *J. Am. Chem. Soc* **2000**, *122*, 11519-11520.

[194] F. E. Romesberg, B. Spiller, P. G. Schultz, R. C. Stevens, *Science* **1998**, *279*, 1929-1933.

[195] A. Heine, E. A. Stura, J. T. Yli-Kauhaluoma, C. Gao, Q. Deng, B. R. Beno, K. N. Houk, K. D. Janda, I. A. Wilson, *Science* **1998**, *279*, 1934-1940.