LBL-36921
UC-408

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

STRUCTURAL BIOLOGY DIVISION

Structural Studies of Bacterial Transcriptional Regulatory Proteins by Multidimensional Heteronuclear NMR

B.F. Volkman
(Ph.D. Thesis)

February 1995

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

DISCLAIMER

LBL-36921
UC-408

Structural Studies of Bacterial Transcriptional Regulatory Proteins by Multidimensional Heteronuclear NMR

Brian Finley Volkman
Ph.D. Thesis

Chemistry Department
University of California, Berkeley

and

Structural Biology Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

February 1995

This work was supported by the Director, Office of Energy Research, Office of Health Effects Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Structural Studies of Bacterial Transcriptional Regulatory Proteins by Multidimensional Heteronuclear NMR

by

Brian F. Volkman

Structural Studies of Bacterial Transcriptional Regulatory Proteins by Multidimensional Heteronuclear NMR

by

Brian Finley Volkman

B.S. (Butler University) 1989

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor David E. Wemmer, Chair
Professor Kenneth Sauer
Professor Sydney Kustu

Abstract

Structural Studies of Bacterial Transcriptional Regulatory Proteins by Multidimensional Heteronuclear NMR

by

Brian Finley Volkman

Doctor of Philosophy in Chemistry

University of California at Berkeley

Professor David E. Wemmer, Chair

Nuclear magnetic resonance spectroscopy was used to elucidate detailed structural information for peptide and protein molecules. A small peptide was designed and synthesized, and its three-dimensional structure was calculated using distance information derived from two-dimensional NMR measurements. The peptide was used to induce antibodies in mice, and the cross-reactivity of the antibodies with a related protein was analyzed with enzyme-linked immunosorbent assays. Two proteins which are involved in regulation of transcription in bacteria were also studied. The ferric uptake regulation (Fur) protein is a metal-dependent repressor which controls iron uptake in bacteria. Two- and three-dimensional NMR techniques, coupled with uniform and selective isotope labeling allowed the nearly complete assignment of the resonances of the metal-binding domain of the Fur protein. NTRC is a transcriptional enhancer binding protein whose N-terminal domain is a 'receiver domain' in the family of 'two-component' regulatory systems. Phosphorylation of the N-terminal domain of NTRC activates the initiation of transcription of genes encoding proteins involved in nitrogen regulation. Three- and four-dimensional

NMR spectroscopy methods have been used to complete the resonance assignments and determine the solution structure of the N-terminal receiver domain of the NTRC protein. Comparison of the solution structure of the NTRC receiver domain with the crystal structures of the homologous protein CheY reveals a very similar fold, with the only significant difference being the position of helix 4 relative to the rest of the protein. The determination of the structure of the NTRC receiver domain is the first step toward understanding a mechanism of signal transduction which is common to many bacterial regulatory systems.

Approved: David A. Dean 12/6/94

Date

Dedicated to my parents

Joe and Barbara Volkman

and to my Grandparents,

who instilled in me

the value of education and

those things greater

Table of Contents

List of Figures

List of Tables

Acknowledgements*

As I near the completion of this summary of my graduate research, I cannot possibly let pass this chance to comment on the environment in which I have spent the last five and one-half years. After finally acclimating myself to Berkeley and all of the ways in which it differs so dramatically from my native Indiana, I have realized that the determining factor in whether this is a 'good' or 'bad' place for me is the people who surround me. In this realization I have begun to despair ever finding another situation which includes so many exceptional individuals. I wish to thank all of those whom I have had the honor of working and playing for making this a very rewarding experience.

In particular, I thank Dave Wemmer for the opportunity to pursue my interests in structural biology in his laboratory. Through the years my admiration and respect for both his intellectual leadership and his dependability have only grown, never diminishing. I hope to continue this relationship long after I have left the lab. I have benefited greatly from interactions with other faculty as well, in particular Loy Volkman, who was at least partially responsible for my decision to come to Berkeley, Joe Neilands and Sydney Kustu, all of whom provided much-needed support for my meager knowledge of the biological sciences. I am also very grateful to David King who has been a very cordial and skilled collaborator at almost every stage of my research.

Almost immediately upon arriving in Berkeley, I became engaged in a close friendship with two other displaced midwesterner first-year chemistry Ph.D. students. Jay Baltisberger, Tony Huff and I shared many triumphs and tragedies as housemates on Neva Ct., surviving the Oakland Hills Fire of 1991 as an adoptive family unit. Within the lab, Mark Kubinec and Marty Pagel also became trusted friends who provided a great deal of the camaraderie of the Wemmer Lab in my first years here.

In more recent times, four people in the lab have been instrumental in any success I have had in my research efforts. Mike Nohaile was instrumental in helping me accelerate

the pace of work on the NTRC project, and has been an excellent companion through the long days and nights of protein purification and NMR assignment. Mark Coy has provided the vast majority of the motivation required to keep the Fur protein project moving, and a Herculean effort in sample preparation. Without the groundbreaking strides of Andrew Lee, the completion of the structure determination of the NTRC receiver domain would not have occurred in my tenure. I expect much greater things in Andrew's future, and will always remember the many hours we spent arguing and laughing, over coffee and lunch, about things scientific and otherwise. Bernie Geierstanger has been an inspiration and I am honored to have had the opportunity to work closely with him on the BD-peptide project. I will continually strive to reach Bernie's level of efficiency and dedication to hard work.

To Ping Lyu, Greg Perens, Jonathon Heller, Fred Damberger, Werner Kremer, Jeff Pelton, Patty Fagan, Rick Storrs, Milton Werner and the other past and present members of the Wemmer group, I am also extremely grateful for the sense of mutual support and friendship which has pervaded the lab. A number of the members of the Pines lab have also provided friendship and NMR kinship, while combining with the Wemmer group to produce the two-time defending champion Magnetic Fielders of the Chemistry Department Softball League (THREE-PEAT!).

Finally, I reserve special thanks for Kristen Iwanaga, who has been an especially caring and supportive companion through the most trying years of my graduate career.

December, 1994

* This work was supported by the Director, Office of Energy Research, Office of Health Effects Research, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Chapter 1 NMR Methodology

Introduction

The determination of protein and peptide structures has been critical to developing a more detailed understanding of the physiology of all biological systems by providing a molecular basis for a wide range of biochemical processes. In addition to the well-established methods of x-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy has become a viable technique for determination of the three-dimensional structure of proteins at atomic resolution. In the last 10 years, development of two dimensional NMR methods (Wüthrich, 1986) and, more recently, three- and four-dimensional heteronuclear techniques (Clore & Gronenborn, 1991) has made structural analysis of small (<150 amino acids) proteins relatively straightforward.

NMR was used to study protein molecules very soon after the advent of pulsed NMR techniques (Wüthrich, 1976). However, it was not until the development of the 2D nuclear Overhauser effect spectroscopy (NOESY) experiment (Kumar et al., 1980) that questions of tertiary structure of proteins could be seriously examined. There are two obstacles to protein structure determination which are overcome by the 2D NOESY experiment.

First, the number of proton (^{1}H) resonances in the NMR spectrum of even a small protein is large, and chemical shift overlap precludes interpretation of the 1D NMR spectrum. Figure 1.1 shows the ^{1}H NMR spectrum of a 124 amino acid protein, illustrating the extreme degeneracy of the resonances. Expanding the spectrum into a second dimension provides a way of reducing this degeneracy by correlating different resonances by way of off-diagonal signals, or crosspeaks. Figure 1.2 shows a contour plot of a 2D NOESY spectrum of the same protein in which the peaks along the diagonal are analogous to the 1D spectrum of Figure 1.1. The signals away from the diagonal represent correlations between protons with different resonance frequencies.

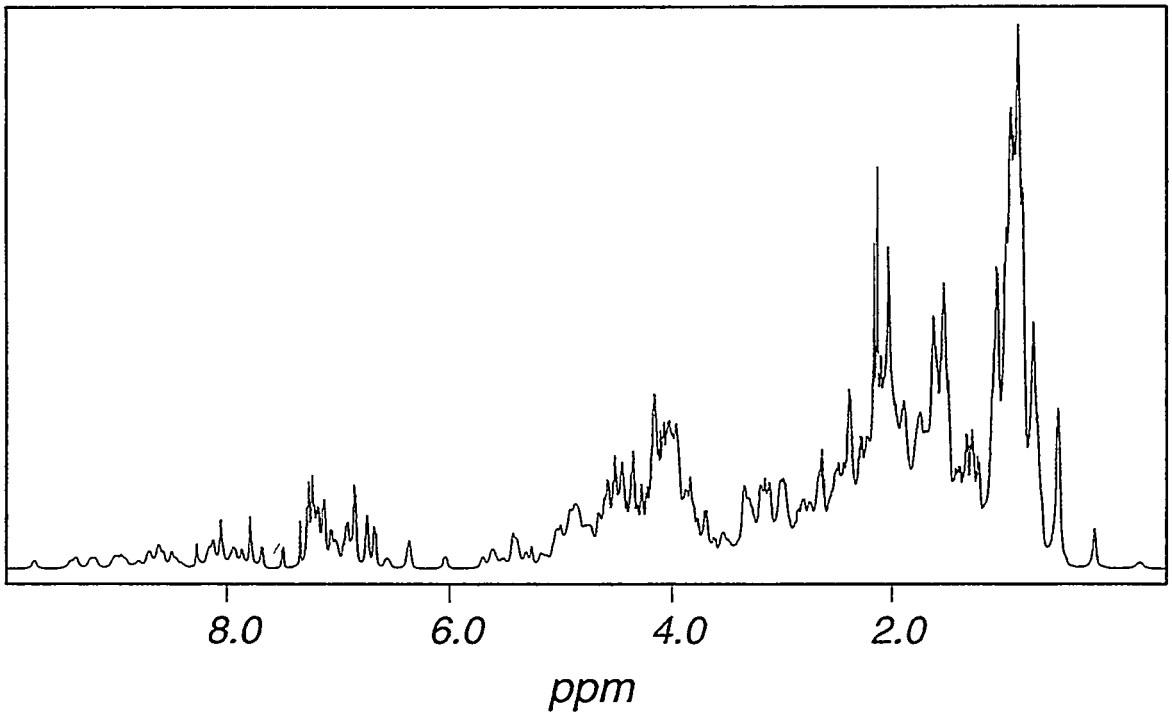


Figure 1.1: ^{1}H NMR spectrum of the NTRC receiver domain at 600 MHz in D_2O, pH 6.4, 25° C. This 124 residue protein has more than 600 separate resonances, and the resulting degeneracy of some proton signals is obvious in the large total intensity of the methyl region (0-2 ppm).

The meaning of these off-diagonal correlations is the second and more significant feature introduced by the NOESY experiment. The nuclear Overhauser effect (NOE) is a result of dipolar relaxation between nuclear spins, which has a $1/r^6$ distance dependence, with r being the distance between a pair of coupled spins. A more complete description of the NOE may be found elsewhere (Noggle & Schirmer, 1971; Ernst et al., 1987). A crosspeak in the NOESY experiment arises due to the NOE between two protons which are within about 5 Å of each other. A number of considerations limit the sensitivity of the NOESY, resulting in the 5 Å cutoff, but because the nature of protein tertiary structure involves close packing of sidechains and repetitive secondary structure elements, a great deal of structural information can be derived from NOE correlations.

By identifying large numbers of NOEs and assigning them unambiguously to specific pairs of protons in the protein, enough information to define the three-dimensional folded structure of a protein is obtained. The information from each NOE crosspeak is converted to the form of a distance restraint, and a procedure for generating a three-dimensional structure of the protein consistent with the set of restraints is followed. The most common methods for structure determination use *distance geometry* calculations (Havel & Wüthrich, 1984) at some stage, often in combination with other types of molecular calculations.

This chapter will present the now-standard multidimensional NMR techniques and structure calculation strategies which are used in the structural analyses described in the following chapters. The description of a few one- and two- dimensional NMR techniques will allow the construction of three- and four- dimensional experimental methods by combination of the one- and two-dimensional 'building blocks'. All pulse programs described are included in Appendix A. A practical approach is taken, emphasizing the application of these experiments to questions of biomolecular structure, while referring to other texts and reviews for a more detailed description of the theoretical basis for these methods.

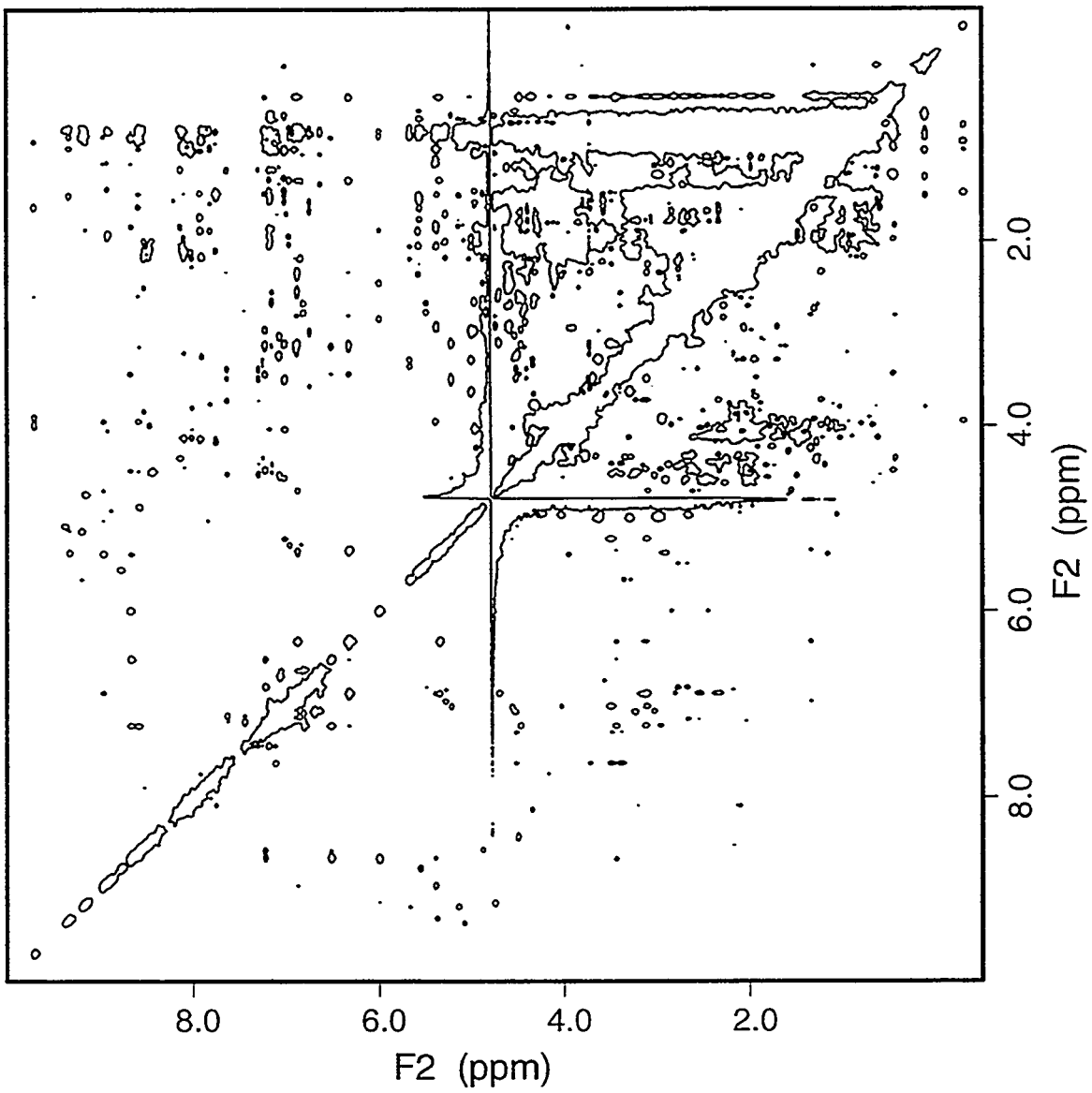


Figure 1.2: 2D ^{1}H NOESY spectrum of the NTRC receiver domain at 600 MHz in D_2O, pH 6.4, 25° C. The NOESY mixing time was 100 ms. Off-diagonal peaks indicate pairs of protons separated by less than 5Å. Overlap of crosspeaks in the upfield region hampers unambiguous assignment of correlated protons, and a similar effect is seen in spectra collected in H_2O solution for the exchangeable amide protons, most of which are not observed in this spectrum.

2D ^{1}H NMR of Proteins

The ^{1}H NMR spectrum shown in Figure 1.1 results from a very simple experimental procedure. In order to acquire this NMR spectrum all that is required is to apply a very short (~10 μs) radiofrequency (RF) pulse at the spectrometer operating frequency (600 MHz). After the RF pulse, the receiver is turned on for ~100-1000 ms and the detected signal (called the *free induction decay*, or FID) is digitized and stored in memory. The time-domain signals in the FID are converted into a frequency domain spectrum by Fourier transformation (FT). The pulse-acquire sequence may be repeated after waiting for the system to move back to equilibrium. This period is called the relaxation delay, and for NMR in liquids will often be on the order of 1-5 s. The one-dimensional spectrum of Figure 1 was acquired in this manner with 128 scans in approximately 5 minutes.

In the simplest example, moving from one-dimensional to two-dimensional (2D) NMR spectroscopy involves simply following the RF pulse with a short delay and another pulse before collecting the FID. The delay is set to zero for the first FID, and then incremented by a fixed value for each successive FID collected. This incremented delay (called *t1*) turns out to be analagous to the direct detection period (called *t2*). The series of FIDs collected with this scheme contains chemical shift information derived during the t2 period as well as information developed during the t1 period. By Fourier transforming the data twice, along both the t2 and t1 dimensions, a 2D frequency spectrum is obtained. By combining pulses and delays (either incremented or constant) an unlimited number of 2D, 3D and 4D NMR experiments can be created.

The off-diagonal signals of 2D NMR experiments contain the useful information, as stated earlier, and not all experiments rely on NOE interactions to produce them. The other type of nuclear interaction which is heavily utilized is the scalar or J-coupling between nuclear spins which are linked by a small number of covalent bonds. The J-coupling

interaction can be described simply as the effect that the state of one spin (+1/2 or −1/2) has on another nearby spin, transmitted by the polarization of the surrounding electron density. J coupling constants are expressed in Hz, and larger values (stronger coupling) generally occur for nuclei which are one or two bonds apart, with J values dropping to zero for nuclei separated by more than three or four bonds. This type of interaction can be either homonuclear (e.g. ^{1}H-^{1}H, ^{13}C-^{13}C) or heteronuclear (e.g. ^{1}H-^{13}C, ^{13}C-^{15}N), and both types are used extensively in the experiments to be described.

In order to use the 2D NOESY experiment to derive distance restraints for structure determination, as outlined in the Introduction, the identity of each resonance must be known in the context of the sequence of amino acids. The process of obtaining sequential assignments requires information from both the 2D NOESY and another 2D experiment, the TOCSY, for TOtal Correlation SpectroscopY (also called HOHAHA, for HOmonuclear HArtman HAhn) (Braunschweiler & Ernst, 1983; Davis & Bax, 1985). The TOCSY provides correlations between groups of protons within each amino acid residue, thereby linking amide, αH, and sidechain protons and allowing the differentiation of residue types. The NOESY experiment contains inter-residue correlations which can then be used to link adjacent residues according to the known amino acid sequence.

2D TOCSY: While the NOESY experiment utilizes the nuclear Overhauser effect (NOE), based upon direct dipolar relaxation of spins, the TOCSY experiment produces correlations based upon a different type of magnetization transfer known as *cross-polarization* in the rotating frame. A complete discussion of the rotating frame may be found elsewhere (Harris, 1986). The TOCSY experiment identifies networks of *J*-coupled spins, also called *spin-systems*. The NH, αH and aliphatic sidechain protons of each amino acid in a polypeptide form a coupled network of spins isolated from the adjacent amino acids by the amide bond, across which all proton *J*-values are essentially zero.

An illustration of the NMR pulse program used for the TOCSY experiment is shown in Figure 1.3B. After a preparation period, a 90° pulse and the t1 delay in which

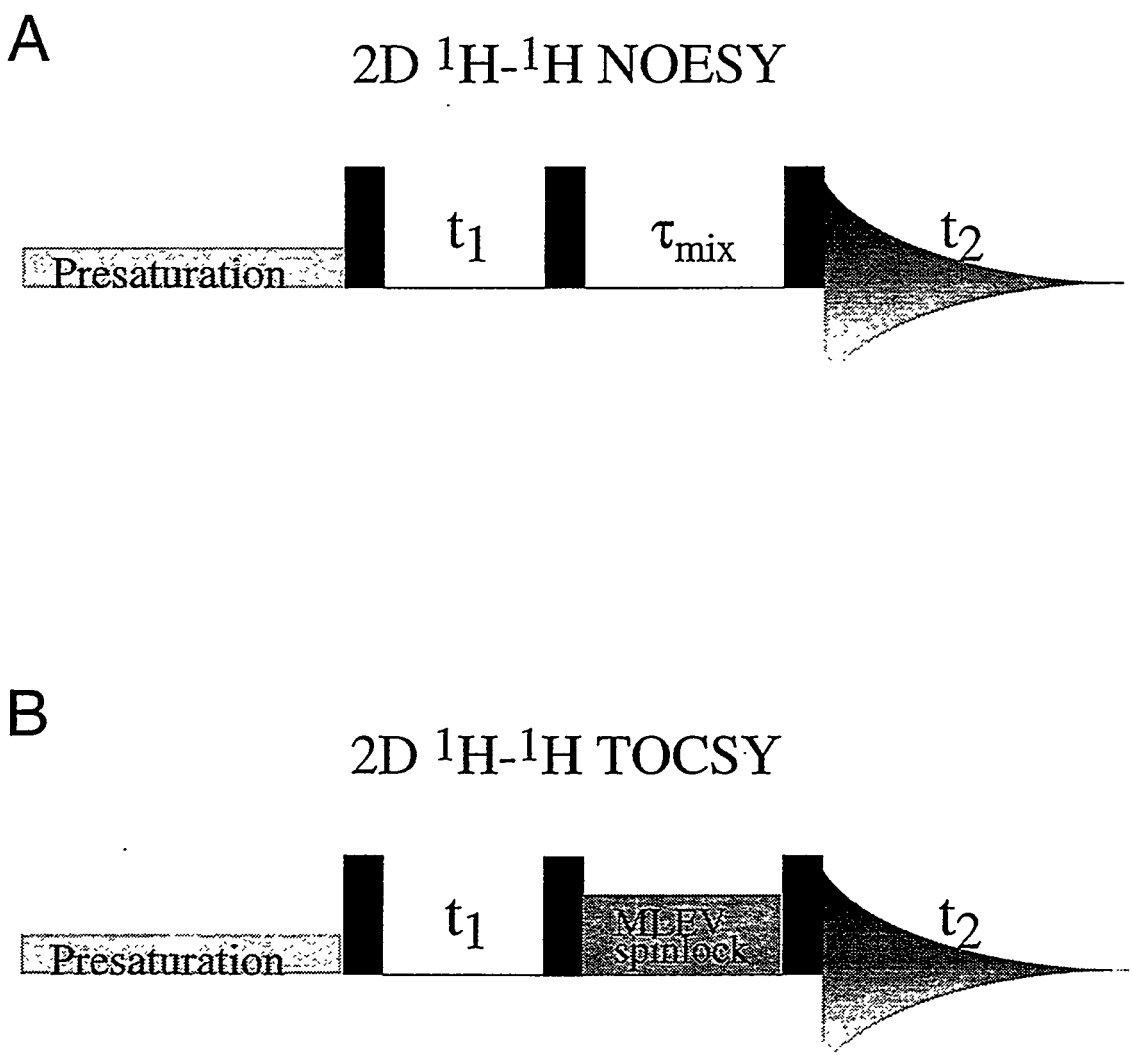


Figure 1.3: Diagrams for 2D ^{1}H NMR pulse sequences. Filled blocks represent 90° high power pulses (typically 10 μs). Presaturation uses long (1-2 s) low power irradiation to selectively saturate the solvent resonance. Indirect evolution (t_1) period is incremented with each block of the experiement. 2D NOESY (A) mixing time may range from 50-450 ms and 2D TOCSY (B) isotropic mixing period uses a repetitive sequence of pulses for spinlock on x- or y-axis, with typical duration 40-100 ms. Pulse programs for AMX-600 NMR spectrometer with complete phase cycles may be found in Appendix A.

chemical shift evolution for the first dimension occurs, a spin-lock field is applied on the *y*-axis. This spin-lock allows coherence transfer via cross-polarization (also called *isotropic mixing*) among all the protons of a spin-system, so that when the magnetization is detected (t2) some spins which evolved under one chemical shift value during t1 evolve under a second chemical shift value during t2. This gives rise to a crosspeak in the 2D frequency spectrum between, for instance, an amide NH resonance at 8.0 ppm and the αH resonance of the same amino acid at 4.6 ppm. With a sufficiently long isotropic mixing period (70-100 ms), it is possible to obtain correlations between the backbone NH and all aliphatic protons of all amino acid types. For a more complete product operator description of the TOCSY experiment, consult the review by Hicks (Hicks et al., 1994).

One technical aspect of ^{1}H TOCSY experiments which is not addressed in the review by Hicks is the reduction of ROESY-type artifacts in TOCSY experiments. The rotating-frame Overhauser effect (ROE) produces the equivalent of NOE information in a spin-lock experiment, closely related to the TOCSY. The impossibility of completely separating the ROESY and TOCSY transfer types has led to the development of the Clean-TOCSY (Griesinger et al., 1988), which incorporates a short NOE mixing period into the TOCSY experiment for the purpose of canceling the ROE, which has the opposite sign of the NOE and TOCSY signals. All of the ^{1}H TOCSY-containing pulse programs in Appendix A incorporate the Clean-TOCSY method.

The intraresidual correlations obtained in the TOCSY experiment must subsequently be assigned to the appropriate position in the amino acid sequence. This is accomplished with the use of the sequential crosspeaks observed in the 2D NOESY experiment.

2D NOESY: The nuclear Overhauser effect (NOE) is used to obtain correlations between spins which are near enough to interact directly via a coupling of the two nuclear dipoles. This coupling involves an exchange of spin states between the two protons, known as *cross-relaxation*, and the rate of this relaxation process has a distance dependence of $1/r^6$, where r is the length of the internuclear vector. The detectable range of this effect is

approximately 5 Å. The 2D NOESY pulse program, shown in Figure 1.3B, is similar to that of the 2D TOCSY, except that the spin-lock isotropic mixing period is replaced with a 90° pulse followed by a delay and another 90° pulse. This delay, labeled τ_m, is the NOESY mixing time. The length of this period can range from 50-450 ms for typical applications in peptides and proteins, and the number and intensity of NOE crosspeaks observed in the 2D spectrum is dependent upon this value. The rate of NOE buildup (i.e., the cross-relaxation rate) is directly related to the molecular weight of the molecule under consideration, so that smaller proteins and peptides may require long mixing times of 200-400 ms which would typically be unsuitable for proteins of greater than 100 amino acids. The population transfer which occurs during the NOESY mixing time gives rise to crosspeaks in the 2D spectrum similar in appearance to the TOCSY experiment. See Figures 2.3, 2.4 and 2.5 for examples of the crosspeaks observed in the two experiments for a small peptide.

Assignment strategy: The partially exclusive nature of the two types of correlations, NOESY and TOCSY, coupled with the regular patterns of NOEs which arise in peptides and proteins, allow the assignment of resonances to sequence-specific protons in the molecule of interest. While both intra- and inter-residue crosspeaks may appear in NOESY spectra, crosspeaks in TOCSY spectra can be due only to intra-residue correlations. Inter-residue NOESY crosspeaks can be used to identify adjacent spin systems in the amino acid sequence. Figure 1.4A shows schematically the types of sequential NOE connections which can be observed. The local geometry of the protons of the polypeptide backbone results in a directionality in the types of NOESY crosspeaks arising between adjacent amino acid residues. For all common values of the dihedral angles ϕ and ψ of a residue *i*, the distance between either the αH or NH of *i* will be less than 5Å from the NH of residue *i+1*, giving rise to a crosspeak in the NOESY spectrum. In many instances, sequential crosspeaks of both types (αH to NH and NH to NH) will be observed for a given *i, i+1* pair, and the common appearance of βH_i - NH_{i+1} crosspeaks

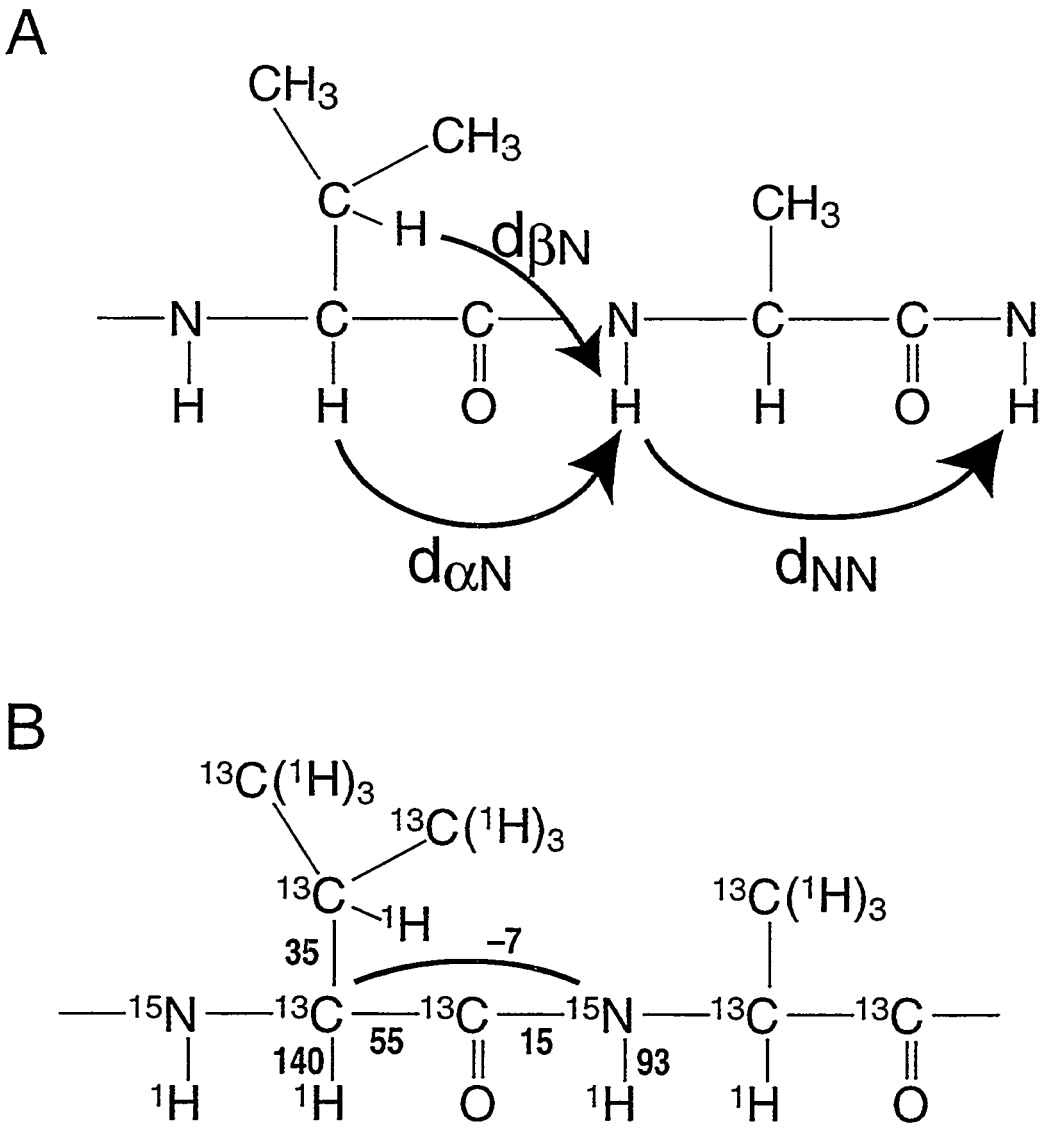


Figure 1.4: Connectivity diagrams for assignment of peptides and proteins. Three types of sequential NOEs (A) allow systematic assignment of resonances in 2D and 3D NOESY spectra by correlation of protons of one residue to the amide proton of the following residue, as indicated by arrows. Heteronuclear *J*-coupling values (B) allow correlation of ^{1}H, ^{13}C and ^{15}N nuclei for both intraresidue and interresidue assignments using 3D and 4D triple-resonance NMR experiments.

provides a total of three possible types of sequential NOESY crosspeaks. This apparent redundancy is often important for resolving ambiguity which may arise due to degeneracy of two or more NH, αH or βH resonances from different spin systems.

In addition to simply providing the sequential context in which to place the resonance assignments of spin systems from the TOCSY experiment, the pattern of intensities of the three types of regular sequential NOEs (abbreviated as $d_{\alpha N}$, $d_{\beta N}$, and d_{NN}) has been found to indicate secondary structural elements. A pattern of strong $d_{\beta N}$ and d_{NN} crosspeaks with weak $d_{\alpha N}$ crosspeaks may indicate helix, while a region with strong $d_{\alpha N}$ crosspeaks and few or no other sequentials is likely to be extended, as in β-sheet structures (Wüthrich, 1986). Other NOE data are required to confirm the presence of secondary structure, and will be discussed further in the context of three-dimensional structure determination.

With the advent of methods for isotopic enrichment of proteins, the application of these basic 2D ^{1}H-only experiments has been superseded by a variety of 2D and 3D multinuclear techniques for obtaining resonance assignments and structural information. In many instances however, the basic sequential assignment strategy described above is still of primary importance, whether utilized in the context of 2D homonuclear or 3D heteronuclear NMR experiments. All of the resonance assignments described in the following chapters have relied primarily on this sequential-backbone NOE-based approach.

Isotopic Labeling

A very significant advance in protein NMR methodology revolves around molecular biology rather than spectroscopy. The development of high-level bacterial overexpression systems for producing large quantities of pure recombinant protein in a defined minimal medium enabled the (relatively) straightforward and inexpensive production of most proteins whose genes had been cloned from the original organism. This allows the incorporation of heteronuclear labels into proteins which would not have been possible for samples isolated directly from tissue of a larger organism. By incorporating a plasmid

DNA encoding the gene of the desired protein into a strain of bacteria (typically an *E. coli* variant) which has been designed to allow precise control of the expression of the desired protein, and then growing liter-scale cultures on medium whose sole sources of carbon and nitrogen are glucose and ammonium chloride (for example), tens to hundreds of milligrams of pure recombinant protein can be isolated with a variety of heteronuclear labeling schemes.

By substituting $^{15}NH_4Cl$ and/or $^{13}C_6$-glucose as the nitrogen or carbon sources, proteins which contain nearly 100% of the desired heteronuclear label can be generated. The usefulness of this approach is derived from the fact that ^{15}N and ^{13}C are NMR-active, spin 1/2 isotopes which can be combined with the ^1H NMR methods already introduced to reduce the degeneracy of the 2D homonuclear experiments (see Figure 1.2) by further increasing the dimensionality. Figure 1.4B shows a dipeptide subunit as a network of NMR-active spins with some of the pertinent *J*-coupling values indicated. The relatively large one-bond *J* values indicated for ^{13}C and ^{15}N nuclei and their directly attached protons were the first heteronuclear couplings to be exploited in 3D approaches to protein NMR, and the basic experiments for detecting these correlations will be described in a later section.

In addition to uniform isotopic enrichment or ^{13}C or ^{15}N, the incorporation of selective heteronuclear labels has also become a very practical approach (McIntosh & Dahlquist, 1990). A particular amino acid or subset of amino acid types can be labeled by growing cell cultures on the same defined medium as used for uniform labeling but with the addition of the isotopically labeled amino acid(s). At the same time, unlabeled amino acids of the other remaining types are added, in order to suppress dilution of the label by transamination or other interconversion mechanisms in the metabolic pathways of the bacterium. In some cases it is desirable to use a mutant strain of bacteria which has one or more enzymes suppressed to reduce the scrambling of a desired labeling scheme.

A reversed approach to selective labeling of proteins has also been recently described in which selectively labeled proteins are obtained by production on minimal medium with $^{15}NH_4Cl$ as the source of heteronuclear labels, and unlabeled amino acids are supplied in various combinations to suppress the synthesis of ^{15}N-labeled amino acids in the cell (Neilands, 1990b). The advantage of this approach lies in the potential reduction in cost, since labeled amino acids are typically more expensive than $^{15}NH_4Cl$.

Yet another type of labeling may become more important as even larger proteins are studied with NMR. Incorporation of deuterium (replacing the spin-1/2 ^1H with spin-1 ^2H) into proteins allows improvement of spectral analysis in two ways (LeMaster, 1990). By generating samples with and without selectively deuterated amino acid types, editing of spectra is possible. This has been useful in studies of large proteins, like Fab fragments of antibodies (Anglister et al., 1984). Alternatively, high levels of random fractional deuteration in proteins can be used to significantly reduce ^1H linewidths by attenuating the contribution of spin-diffusion to overall relaxation times. As pulse progams become longer with increasing dimensionality and number of magnetization transfer steps and relaxation times grow shorter with increasing protein size, the use of deuterium labeling may become a requirement for NMR assignment strategy.

At this time 1g of $^{15}NH_4Cl$ costs approximately $50, enough for a 1L bacterial cell culture, compared to approximately $100-$750 for the amounts of various labeled amino acids needed for a 1L culture (typically 100-200 mg). Because $^{13}C_6$-glucose is more expensive (~$600/g) and it is used at the rate of 2 g/L of cell culture, ^{13}C labeling requires a larger investment in samples, and demands careful optimization of the expression and purification of the protein of interest. Typically, acceptable levels of protein production are in the 10-50 mg/L range. Likewise, ^{13}C-labeled amino acids are more costly, and selective enrichment techniques involving ^{13}C-labeled amino acids are therefore less commonly used. In general, to produce the combination of uniformly and selectively labeled samples needed for efficient assignment of resonances and structure determination of proteins in the

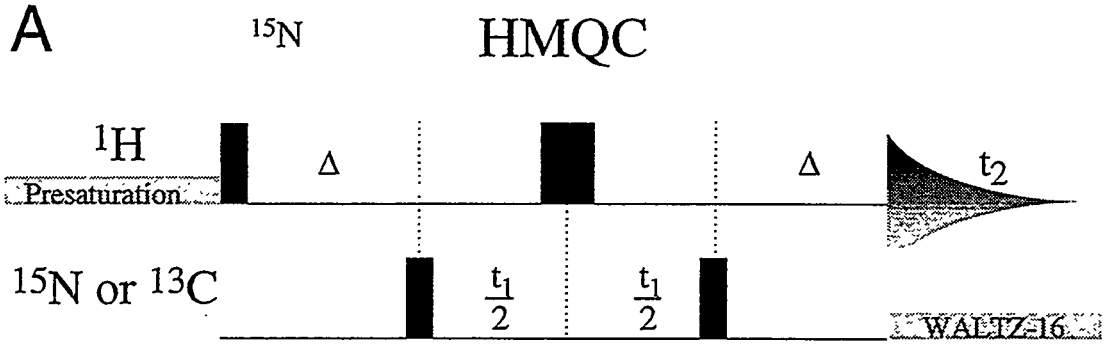
10-20 kD range, it can be expected that $1,000-$2,000 worth of isotopes or more will be required. In the next section the basic experiments of heteronuclear NMR will be described, and the relationship between the information obtained from both selective and uniform labeling strategies will be addressed.

2D Heteronuclear Correlation Experiments

The basic scheme for detecting the ^{15}N or ^{13}C nuclei which have been incorporated into a peptide or protein is a 2D correlation experiment, of which there are two distinct types: Heteronuclear Multiple Quantum Coherence (HMQC) and Heteronuclear Single Quantum Coherence (HSQC). These experiments will be described in some detail, since when combined with the NOESY and TOCSY, they allow the generation of a nearly unlimited number of 3D and 4D experiments for structural protein studies. In each case the correlations observed are generated by polarization transfer between the heteronucleus and the directly attatched proton. This transfer is mediated by the large one-bond couplings between these pairs of nuclei ($J^1_{NH} \approx 93$ Hz, $J^1_{CH} \approx 140$ Hz).

Heteronuclear Multiple Quantum Coherence (HMQC): The HMQC experiment is diagrammed in Figure 1.5A. In addition to a schematic of the pulse program, a simplified description of the product operator evolution during the course of the experiment is given in Figure 1.5B. As the name implies, the correlation between proton and heteronucleus is generated in the form of a multiple quantum operator ($I_y S_x$), which is allowed to evolve under only the heteronuclear chemical shift Hamiltonian operator during the indirect detection period, t1. The Δ periods serve to convert in-phase proton magnetization into antiphase magnetization (and back again) via evolution under the heteronuclear scalar coupling Hamiltonian. This deay is ideally set to $1/(2J)$, which in the case of amide ^{15}N-^1H pair is ~5.3ms. For proteins a value of 4.6ms is often used to reduce the effects of relaxation, which causes the exponential decay of the total magnetization.

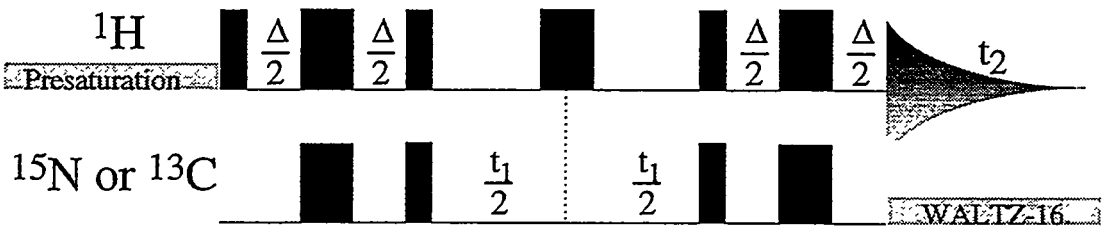
A



B

$$I_z \xrightarrow{90^\circ} I_x \xrightarrow{\Delta} I_y S_z \xrightarrow{90^\circ} I_y S_x \xrightarrow{t_1} I_y S_y \times \xrightarrow{90^\circ} I_y S_z \xrightarrow{\Delta} I_x(\sin \omega_s t_1)$$
$$(\sin \omega_s t_1)$$

C



D

$$I_z \xrightarrow{90^\circ} I_x \xrightarrow{\Delta} I_y S_z \xrightarrow[90^\circ]{90^\circ} I_z S_x \xrightarrow{t_1} I_z S_y \times \xrightarrow[90^\circ]{90^\circ} I_y S_z \xrightarrow{\Delta} I_x(\sin \omega_s t_1)$$
$$(\sin \omega_s t_1)$$

Figure 1.5: Diagrams for 2D HMQC (A) and HSQC (C) pulse sequences. Narrow and wide filled blocks represent 90° and 180° high power pulses, respectively. Delays Δ are set to approximately 1/(2J), where J is the one-bond H-X coupling constant. Product operator descriptions of the HMQC (B) and HSQC (D) experiments indicate the presence of multiple quantum coherence in the HMQC and antiphase single quantum coherence in the HSQC. Coefficients of the density matrix are omitted, except for the inclusion of chemical shift evolution during t_1.

The product of this experiment is a 2D spectrum in which the indirect (F1) dimension contains the heteronuclear chemical shift value, and the directly detected (F2) dimension records the chemical shift of the attached proton. Figure 3.2B shows the HMQC spectrum of a uniformly ^{15}N-labeled protein. The advantage that a heteronuclear dimension provides is immediately obvious. While the degeneracy of resonances in the 1D NMR spectrum of a protein is partially resolved in the 2D NOESY experiment (Figure 1.2), extreme degeneracy of amide NH resonances is still a problem in the homonuclear case. The 2D HMQC provides a two-dimensional 'fingerprint' of the protein in which nearly every amide is uniquely identified by two chemical shifts, ^{15}N and ^{1}H, rather than the single proton value observed in the homonuclear case. The importance of obtaining unique chemical shift values for the backbone amides is clear from the previous description of assignment methods, since all of the sequential NOE patterns rely on at least one amide proton chemical shift.

The usefulness of selective labeling is most obvious in the examination of HMQC spectra collected on both selectively labeled and uniformly labeled proteins. Figures 3.2B, 3.5 and 3.6 show the HMQC spectra of the Fur(77-147) protein with uniform-^{15}N, ^{15}N-Leu, and ^{15}N-His incorporation. Comparison of these spectra, combined with 3D spectra collected on the uniformly ^{15}N-labeled sample can simplify the assignment process by immediately identifying the amino acid type of a particular spin system. This has been especially useful in studies of the Fur(77-147) protein (see Chapter 3), where short relaxation times prevent the acquisition of high quality TOCSY spectra. When 3D experiments are discussed further, it will also be apparent that selective labeling can be used to simplify 2D NOESY and TOCSY spectra by incorporating an HMQC-like filter into the pulse program and using the heteronucleus to edit spin systems.

Heteronuclear Single Quantum Coherence (HSQC): In addition the the HMQC, a closely related but distinct experiment is also used to extract the information contained in labeled protein samples. The HSQC experiment is diagrammed in Figure 1.5C, and an

HSQC spectrum is shown in Figure 4.1. It is apparent that the pulse sequence for the HSQC is more complex than the HMQC, however the data obtained from both experiments are essentially the same. The fundamental difference is in the specific form of the density matrix during the ^{15}N evolution (t1) period, shown in Figure 1.5D. As stated above, in the HMQC the density matrix is comprised of a combination of zero- and double-quantum coherences (also called multiple-quantum coherence); in contrast, the $I_z S_x$ operator which is present during t1 in the HSQC is a single quantum coherence on the ^{15}N nucleus, hence the name Heteronuclear Single Quantum Coherence (Bodenhausen & Ruben, 1980). The primary advantages of this approach are the favorable relaxation characteristics of the single quantum operator versus the multiple quantum operator, and reduced splitting of resonances in indirect (t1) dimension.

Because the multiple quantum operator has a ^{1}H component in the transverse plane, it is subjected to the faster T_2 relaxation rates of ^{1}H nuclei, while the single quantum operator in the HSQC contains only longitudinal ^{1}H magnetization, and decays more slowly. The net effect is to produce potentially higher sensitivity and demonstrably higher resolution in the HSQC compared with a identically acquired HMQC experiment, if both are collected with high resolution in the t1 dimension (Norwood et al., 1990). This distinction is not always apparent in 3D and 4D experiments, where resolution in the indirect dimensions is generally digitization limited, rather than linewidth limited. For this reason many of these experiments still use the HMQC scheme rather than HSQC. In addition to these basic descriptions of the two common types of heteronuclear correlation experiments, a few more specific aspects and variations will be discussed.

Binomial Water Suppression HMQC: Because it has been essential to collecting high-quality HMQC spectra on some protein samples, it is worth mentioning a specific version of the HMQC with an alternative method of solvent suppression. The '1-1 echo' scheme for selective excitation was developed (Sklenar & Bax, 1987) as an improvement upon the basic binomial solvent suppression methods, the simplest of which used a single

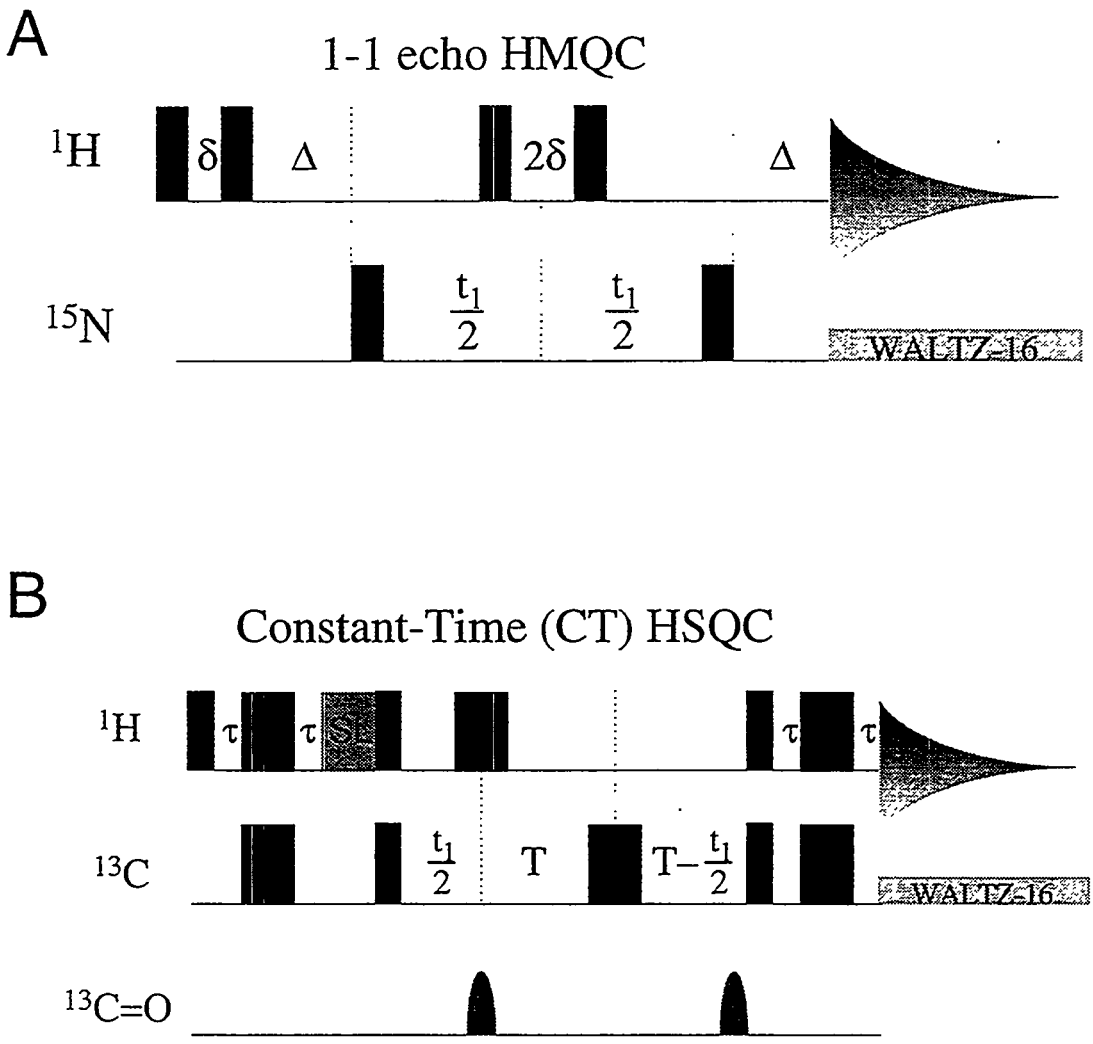


Figure 1.6: Diagrams for '1-1 echo' HMQC (HMQC-JR) (A) and CT-HSQC (B) pulse sequences. A. The HMQC-JR (for jump-return) experiment is often used with protein samples which have increased amide exchange rates due to elevated pH. Delays δ and 2δ are required for water suppression, and determine the excitation profile for the ^{1}H dimension. B. The CT- (constant time) HSQC derives chemical shift evolution in t_1 during a constant delay (2T), across which a ^{13}C 180° pulse is shifted. The signal obtained in the t_1 dimension has no decay due to relaxation, which results in much higher resolution in the frequency domain. Magnetization transfer during Δ delays occurs as in normal HMQC and HSQC.

pair of 45° or 90° pulses separated by a short delay, referred to as the '1-1' or 'jump-and-return' technique. The 1-1 echo is incorporated into the HMQC experiment by simply replacing each of the two ^{1}H pulses by a pair of 90° pulses separated by delays, as shown in Figure 1.6A. The first pair is separated by the delay used in a normal 1-1 excitation scheme. The value of this delay, δ, is determined by the desired excitation profile, and is set to $1/(2\Delta)$, where Δ is the distance in Hz from the carrier frequency to the next null in the excitation profile. In the 1-1 echo HMQC, the second pair of 90° pulses is separated by 2δ in order to achieve the refocusing of the ^{1}H evolution during t1. The resulting experiment achieves extremely high levels of solvent suppression, has a flat baseline in t2, and requires the use of no presaturation. This is a great advantage in instances where NH exchange rates are high due to pH or lack of secondary structure, or when spin diffusion causes saturation transfer during preirradiation.

Constant-Time HSQC: A variation on the HSQC which has been incorporated into many recently developed triple-resonance (^{1}H, ^{15}N, and ^{13}C) experiments is the constant-time (CT) HSQC (Vuister & Bax, 1992). In this experiment, total evolution period, which is normally incremented by a fixed amount with each block of a 2D experiment, is fixed at a constant value. A diagram of a ^{13}C CT-HSQC pulse program is shown in Figure 1.6B Chemical shift evolution of the indirectly detected nuclei is obtained by shifting the position of a ^{13}C 180° pulse across the CT delay period as t1 increases. The ^{13}C 180° pulse does not affect ^{13}C-^{13}C scalar couplings, and since the 2T period is fixed, no homonuclear splittings appear in the ^{13}C spectrum. The value of 2T is adjusted to maximize the observable magnetiztion at the end of the delay (usually a multiple of 1/J). The advantage of the CT method lies in the high resolution which may be obtained, due to the lack of decay and absence of homonuclear splittings observed in the indirect dimension. As an example of this high resolution, a region of the ^{13}C CT-HSQC spectrum of the NTRC receiver domain is shown in Figure 1.7. The primary disadvantage is the reduced sensitivity due to the long delay imposed on the early t1 points, which in a non-CT

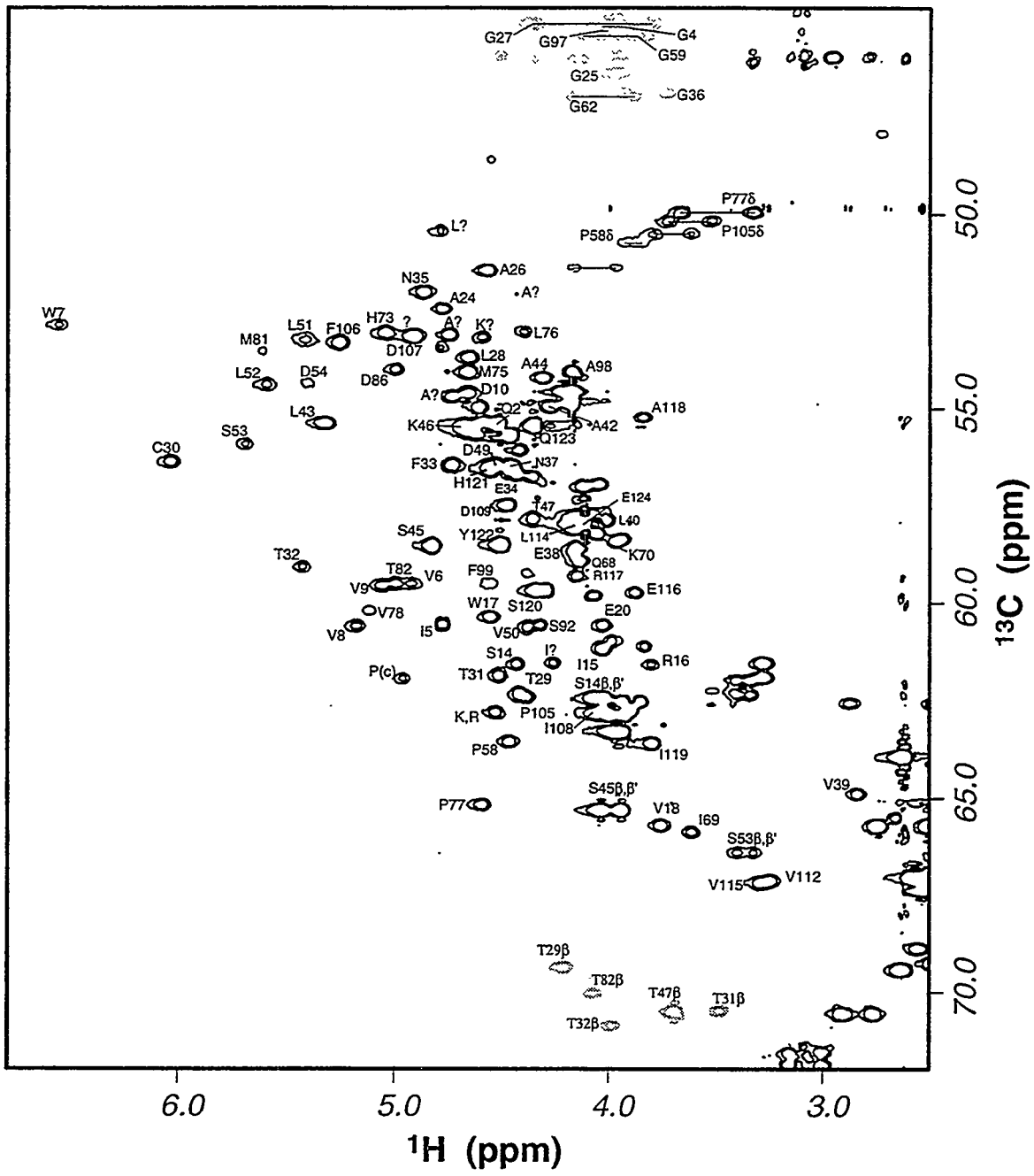


Figure 1.7: A portion of the ^{1}H-^{13}C CT-HSQC spectrum of the NTRC receiver domain. Positive contours are shown in black, and negative contours are shown in gray. The α region is shown; also included are threonine β's and proline δ's. Because the CT delay, 2T, is set to $1/J_{CC}$, carbons with an odd number of attached carbon atoms give rise to peaks of opposite sign from those attached to an even number of carbons.

experiment would not be subject to the magnetization loss due to relaxation during the CT period. However, in many ^{13}C experiments involving protein samples which are uniformly labeled at nearly 100%, this is the only way to obtain high-resolution, carbon-decoupled spectra.

INEPT: In some of the experiments to be described below, a 3D experiment will be created simply by concatenating two of the 2D experiments described above; in other cases the description is not quite so simple. In every case, however, the experiment can be broken down into relatively simple steps called *polarization transfers*. In most cases, the polarization transfer steps are simply the first half of the basic HSQC experiment described above. This type of transfer was actually described before the advent of the 2D heteronuclear correlation experiments, and was termed *I*nsensitive *N*uclear *E*nhancement via *P*olarization *T*ransfer, or INEPT (Morris & Freeman, 1979; Burum & Ernst, 1980). This experiment was initially designed to increase the sensitivity of direct detection of insensitive nuclei like ^{15}N or ^{13}C by transfering the relatively high levels of polarization of the attached ^{1}H nucleus, using the large one-bond *J*-coupling. In the HSQC, an INEPT transfer is performed, the t1 period allows chemical shift evolution on the heteronucleus, and then a *reverse*-INEPT transfers the polarization back to the ^{1}H for direct detection.

Current 3D and 4D triple-resonance experiments use many INEPT polarization transfer steps to move the magnetization from ^{1}H to ^{13}C to ^{15}N, and eventually back to ^{1}H, although the first triple-resonance 3D experiments used both HMQC and INEPT transfers (Kay et al., 1990). Of constant concern when designing experiments with multiple polarization transfer steps are the effects of relaxation, and the reduction in sensitivity each time an additional delay is incorporated into the pulse program. To address this problem, schemes for performing chemical shift evolution and polarization transfer concomitantly have been devised, as a way to shorten the total time required to complete the pulse train and gain sensitivity. Although the basic components of nearly all multidimensional heteronuclear NMR experiments have been mentioned here, the continual refinement of the

most useful approaches leads to pulse programs which are increasingly difficult to decipher in terms of these basic building blocks.

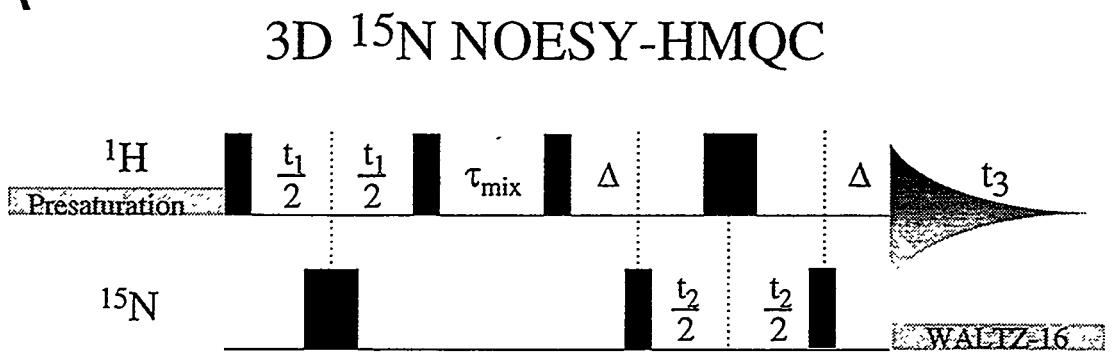
Three-Dimensional ^{15}N-separated Experiments

With the basic two-dimensional experiments as building blocks, the first basic heteronuclear 3D NMR experiments can be easily described. Starting with either the 2D NOESY or TOCSY experiment shown in Figure 1.3, and appending it with the ^{15}N HMQC (Figure 1.5A) results in either a 3D NOESY-HMQC or 3D TOCSY-HMQC (Kay et al., 1989; Marion et al., 1989b). These pulse programs are illustrated in Figure 1.8. These are experiments which can be performed on samples which are labeled with only ^{15}N, and which can be used successfully for the sequential assignment process while also providing a large amount of information on secondary structure in favorable cases.

These three-dimensional experiments produce a dataset in which the densely packed information contained in the 2D NOESY and TOCSY has been expanded into a vertical dimension, resulting in a series of sparsely populated 2D ^{1}H-^{1}H planes. Figure 1.9 shows a comparison of the 2D NOESY and a plane from the 3D ^{15}N-separated NOESY-HMQC experiments, collected on the 14kDa N-terminal domain of the NTRC protein. Clearly, spin systems which were highly degenerate in the NH region of the 2D NOESY have been isolated by the chemical shift of the amide ^{15}N. The 3D NOESY- and TOCSY-HMQC datasets can be analyzed in a similar fashion to that described above for the 2D NOESY and TOCSY. A convenient way to display the information obtained in the assignment process is in the form of strips corresponding to spin systems, collected into a single plane and arranged side-by-side in order of the sequence. An example of this type of 'strip plot' is shown in Figure 4.2, and the related discussion in Chapter 4 describes the types of NOEs observed and their relationship to secondary structure determination.

The usefulness of these experiments is typically limited by the observed linewidths of the protein to be studied. This effect is manifested in two ways. First, because the

A



B

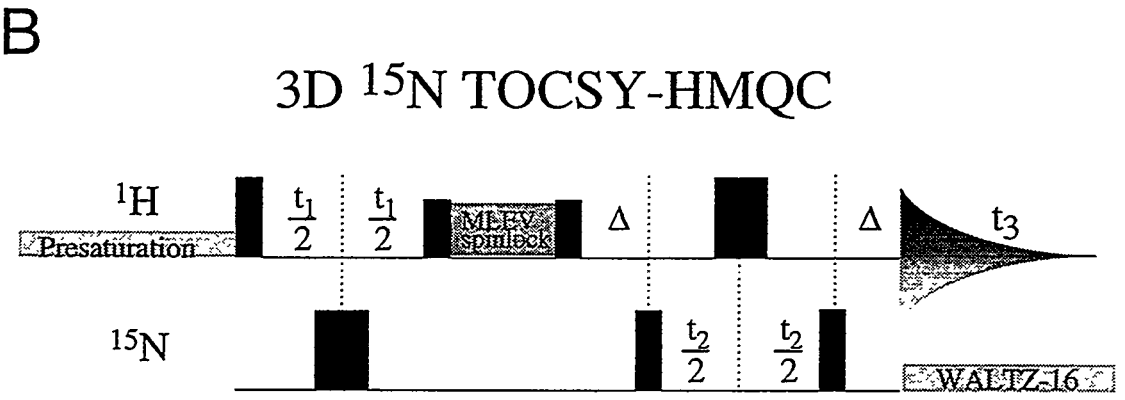


Figure 1.8: Diagrams for 3D ^{15}N NOESY-HMQC (A) and 3D ^{15}N TOCSY-HMQC (B) pulse sequences. Both experiments are derived from the combination of the 2D NOESY, TOCSY, and HMQC pulse sequences shown previously. Typical application of these pulse programs involves the aquisition of 64 FIDs in the t_2 (^{15}N) dimension, 256 FIDs in the t_1 (^{1}H) dimension, with 8 or 16 scans per FID. Total experiment time for this scheme is approximately 3 days. Analogous experiments which use HSQC instead of HMQC have also been developed (see Appendix A).

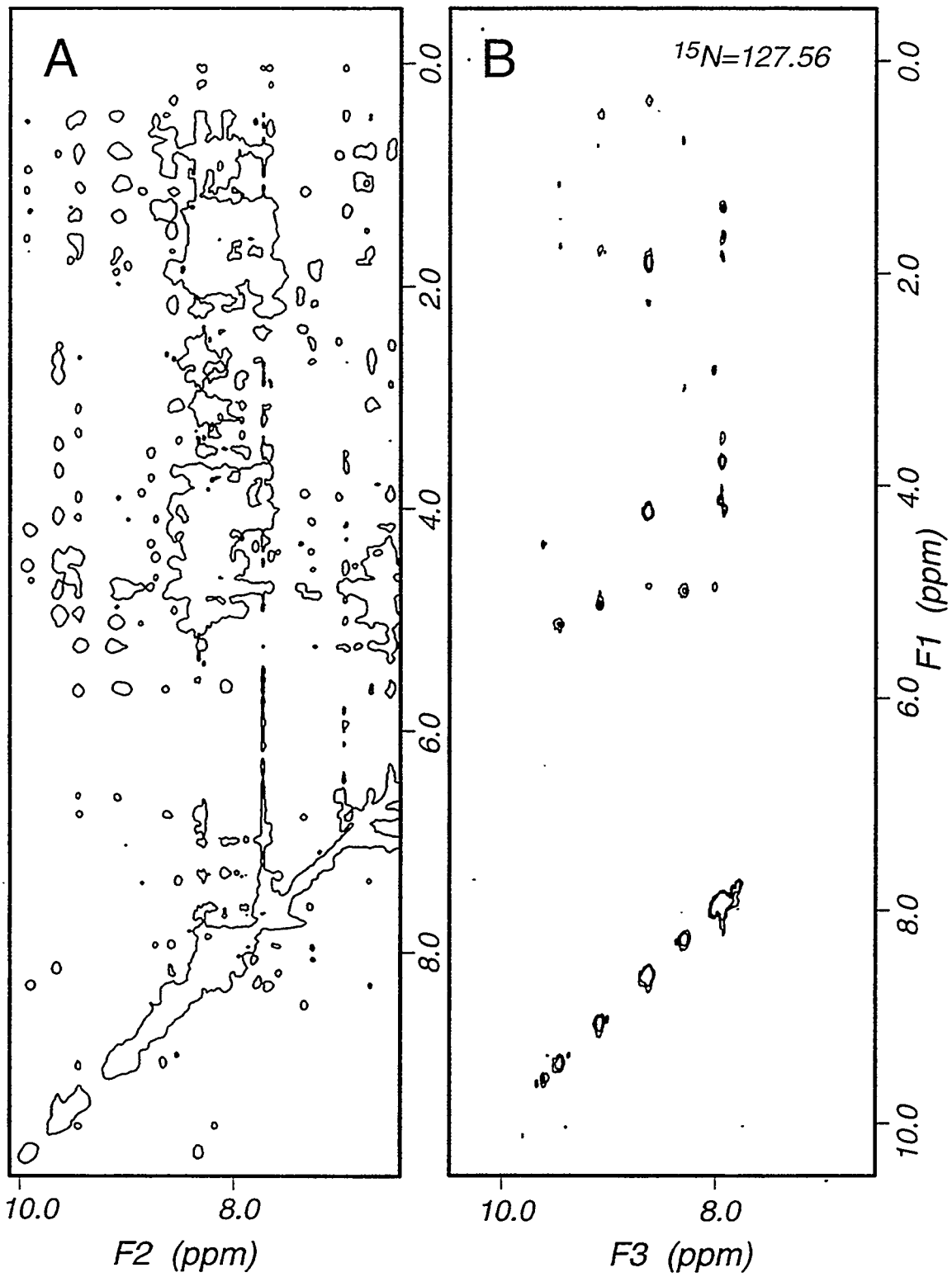


Figure 1.9: Comparison between 2D NOESY (A) and a single plane of a 3D ^{15}N NOESY-HMQC (B), both collected on the FurT88 protein. By separating spin systems in a third dimension (64 ^{1}H-^{1}H planes) according to ^{15}N chemical shift, crosspeaks can be unambiguously correlated to a single NH, as shown in panel B.

number of FIDs to be collected in a 3D experiment is large, the number of scans per FID must be limited, typically to 8 or 16. For proteins with broad lines, this may not provide sufficient signal intensity. This lack of signal is tied to the increased loss of magnetization due to relaxation during the pulse sequence. This is especially noticeable in the 3D TOCSY-HMQC experiment, where signal loss during the isotropic mixing period can be dramatic.

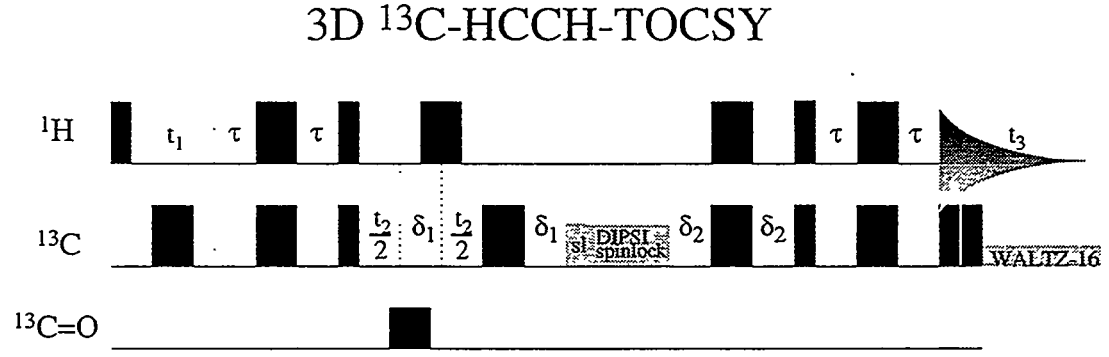
3D ^{13}C-Separated and ^{13}C,^{15}N, ^{1}H Triple Resonance Experiments

The key to overcoming the problems associated with ^{1}H-^{1}H TOCSY transfer in proteins with broad lines lies in the ability to uniformly incorporate ^{13}C into the sample of interest. As indicated in Figure 1.4B, uniform ^{13}C and ^{15}N labeling results in a network of spins with one-bond J-coupling values significantly larger than the small (<10 Hz) three-bond J-values which the ^{1}H-^{1}H TOCSY relies upon for correlations. This network extends across the peptide bond as well, opening the door to non-NOE-based approaches to the task of obtaining sequential assignments. Two experiments which rely on this dense network of J-coupled spins have been used for protein assignments in this work, and they will be described below.

3D ^{13}C HCCH-TOCSY: The most efficient experiment for obtaining complete sidechain assignment information, HCCH-TOCSY (Bax et al., 1990), utilizes isotropic mixing of the aliphatic ^{13}C spins to achieve total correlation of sidechain resonances. The pulse sequence for this experiment is shown schematically in Figure 1.10A. A description of this experiment is fairly straightforward in terms of the preceding discussion of 2D techniques.

First, ^{1}H chemical shift evolution occurs in the incremented t1 period, just as in the 2D ^{1}H NOESY and TOCSY. An INEPT then transfers magnetization to the directly attached ^{13}C nuclei, which then undergoes ^{13}C chemical shift evolution in t2. Additional delays which refocus the antiphase ^{13}C magnetization are essential to producing the in-

A



B

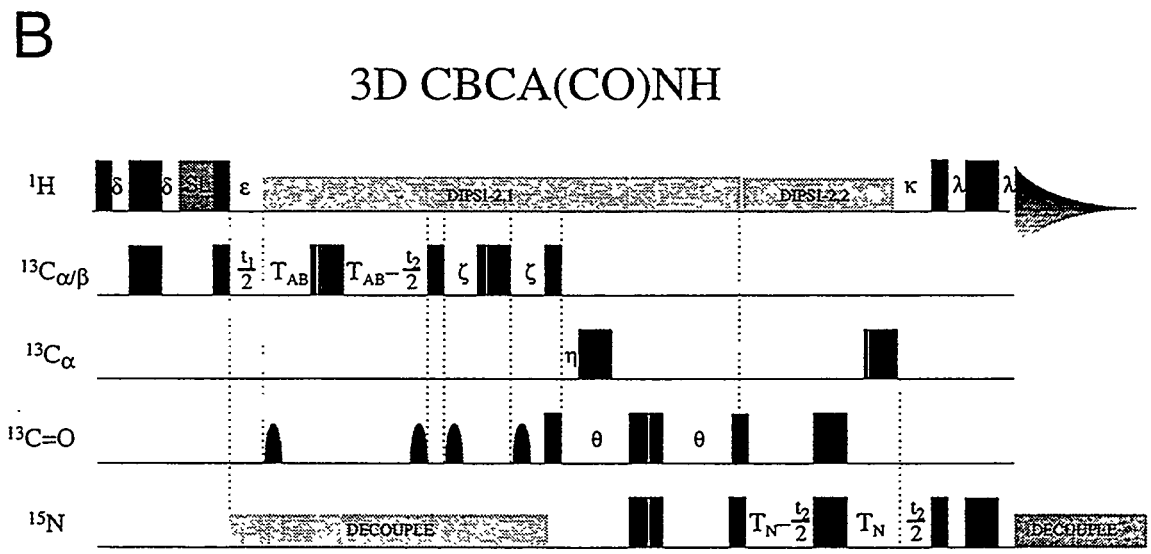


Figure 1.10: Diagrams for 3D ^{13}C HCCH-TOCSY (A) and 3D CBCA(CO)NH (B) pulse sequences. A. 3D HCCH-TOCSY uses a 25 ms duration DIPSI spinlock field for isotropic mixing of ^{13}C magnetization; τ is set to $1/(4J_{CH})$ (1.5 ms) for INEPT transfers, while δ_1 and δ_2 are set to $1/(6J_{CH})$ (1.1 ms) to allow equal refocusing of methyl, methylene, and methine CH magnetization. B. The CBCA(CO)NH uses a CT t_1 period of 6.6 ms for ^{13}C chemical shift evolution and CT t_2 period of 22.2 ms for ^{15}N evolution. Magnetization is transferred from ^{1}H to ^{13}C during the δ delays (1.5 ms), and from ^{15}N back to ^{1}H during λ delays (2.25 ms). Other delays (ε = 2.3 ms, ζ = 3.7 ms, η = 4.5 ms, θ = 11.4 ms, κ = 5.4 ms) are used for refocusing antiphase terms after INEPT transfers or transfer steps throught the carbonyl ^{13}C nucleus.

phase signals which undergo isotropic mixing during the DIPSI (Shaka et al., 1988) spinlock period. A reverse INEPT transfers the ^{13}C magnetization back to aliphatic protons for detection in t3. One 180° pulse on ^{13}C=O nuclei in the middle of t2 is required to prevent dephasing of C_α spins due to the one-bond coupling.

The resulting 3D dataset contains a highly resolved and redundant set of correlations which allow the complete assignment of most ^{1}H and ^{13}C resonances of a protein. One plane of the HCCH-TOCSY spectrum of the NTRC receiver domain is shown in Figure 4.3. A number of different spin systems are observed in that plane, and it is important to note that four different ^{13}C chemical shifts are represented by a single plane in the F2 dimension. This apparent ambiguity is a designed attempt to achieve higher resolution in the ^{13}C dimension through the use of *aliasing*.

Because the range of ^{13}C chemical shifts is so large, it is difficult to obtain high-resolution data without an inordinately high level of digitization in the ^{13}C dimension. The strong general correlation between ^{13}C and ^{1}H chemical shift values for the different types of CH pairs (α, β, γ, etc.) makes it possible to narrow the spectral width of an indirectly detected ^{13}C dimension, in the process locating, or aliasing, more than one actual ^{13}C shift at a given position in the spectrum. It is essential to first examine a high-resolution CT-HSQC which has been collected with a wide spectral width before determining the appropriate spectral width for the HCCH-TOCSY. As long as care is taken in designing the experimental parameters, the true ^{13}C chemical shift may be extrapolated on the basis of the associated ^{1}H chemical shift.

In the case of work on the NTRC receiver domain, the combination of the HCCH-TOCSY and the 3D ^{15}N NOESY-HMQC and TOCSY-HMQC experiments was sufficient to allow assignment of most of the resonances of the protein. Another experiment was collected, however, which very helpful in confirming assignments by providing correlations between $^{13}C_\alpha$ and $^{13}C_\beta$ resonances and the amide ^{1}H and ^{15}N resonances of the adjacent residue.

CBCA(CO)NH: The only triple-resonance experiment used in this work provides sequential assignment information which would normally be used in combination with a related experiment, such as the CBCANH, designed to give the complementary intra-residual information (Grzesiek & Bax, 1992b). The CBCA(CO)NH (Grzesiek & Bax, 1992a) uses a series of INEPT transfers to connect the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ nuclei of one residue with the NH group of the following residue, via the series of scalar couplings between the $^{13}C_{\alpha}$, $^{13}C_{\beta}$, C' and ^{15}N nuclei. Constant-time evolution periods record chemical shift information from the $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and ^{15}N nuclei as magnetization is passed to them. A diagram of the pulse program is shown in Figure 1.10B.

In Figure 1.11 a series of plots extracted from the CBCA(CO)NH spectrum of the NTRC receiver domain illustrate the simplicity of the data obtained. Because of the relatively high signal-to-noise ratios and sparseness of this and similar experiments, they have been examined as the most promising basis for automated assignment methods by a number of researchers. In some cases, programs have been designed which have succeeded in performing automated sequential assignment of a protein with extremely high quality spectra, but a widely applicable method is not yet availible. Perhaps the most obvious drawback to this and the related experiments is the reliance upon detection of the amide protons, which can be attenuated by exchange with solvent in systems at near-neutral pH. This problem is evident in all of the experiments described, and it is hoped that the development of experiments using pulsed-field gradients for solvent suppression will reduce or eliminate the difficulty.

4D $^{13}C/^{13}C$-edited HMQC-NOESY-HMQC: The final NMR experiment to be described is of the highest dimensionality, which belies the relatively simple pulse sequence, shown in Figure 1.12A. The 4D $^{13}C/^{13}C$-edited NOESY (Clore et al., 1991) is an essential component in the development of information on three-dimensional structure of a protein. It allows the unambiguous assignment of a large number of NOEs between sidechain protons which would normally be highly degenerate in the 2D NOESY shown in

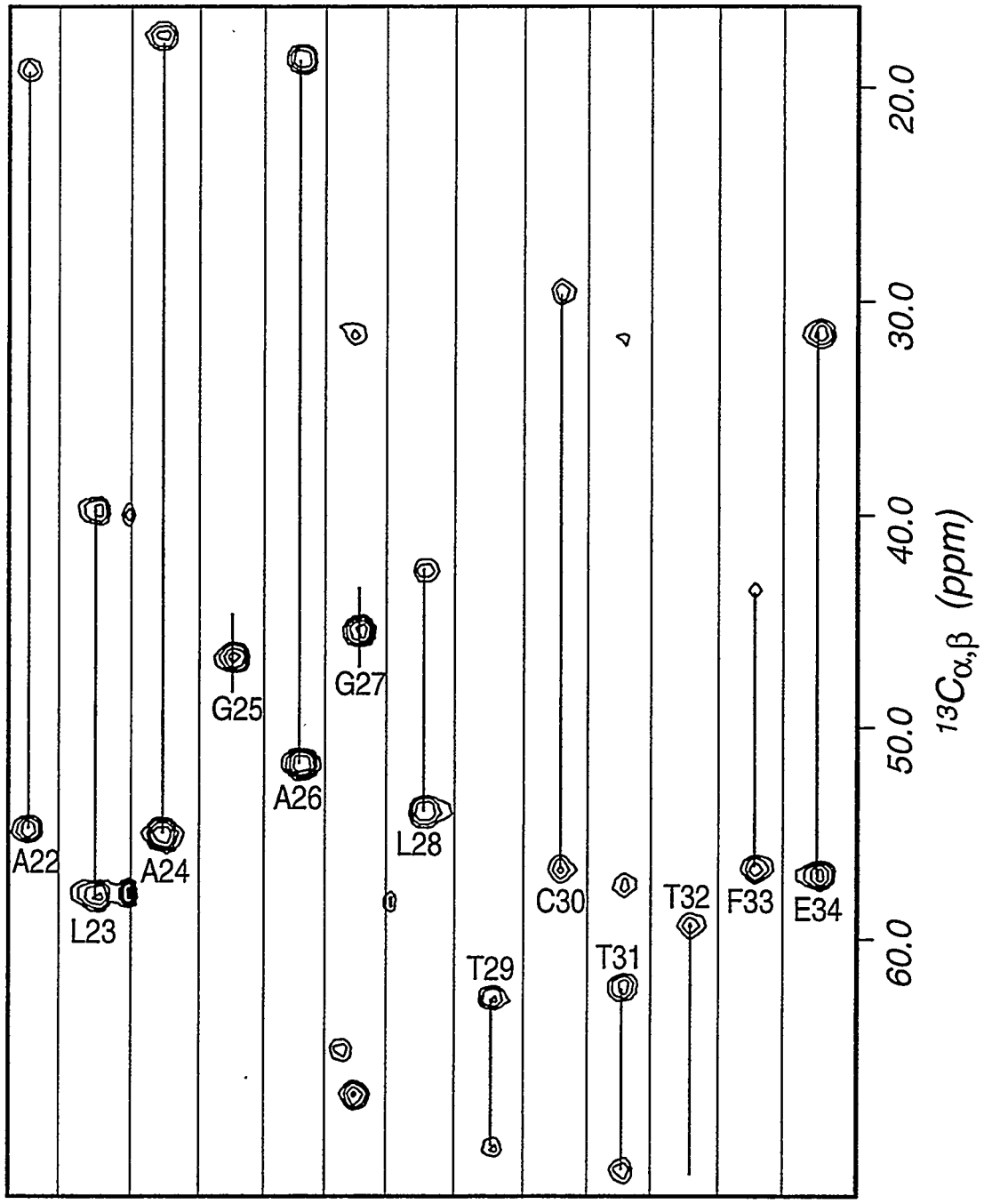


Figure 1.11: Strip plot of selected regions from CBCA(CO)NH spectrum of the NTRC receiver domain. Peaks correlate $^{13}C(\alpha, \beta)$ chemical shifts with the ^{15}N and ^{1}H chemical shifts of the following residue. The $^{13}C_{\alpha}$ correlations are labeled with residue type and number, and vertical lines connect $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ correlations. Strips are taken from the plane corresponding to the amide ^{15}N chemical shift of the following residue, and are centered at that amide ^{1}H chemical shift. $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts alone can be useful in identifying amino acid type; comparison of the patterns of the threonine, glycine and alanine residues provides a good example of this.

Figure 1.2, and resolves much of the overlap which can make interpretation of 3D ^{13}C-edited NOESY data difficult.

The 4D $^{13}C/^{13}C$-edited NOESY as implemented in this lab by Andrew Lee follows the original description by Clore et al. very closely. A minimum of 8 scans per FID are required to complete the phase cycle, which has been designed to minimize artifacts. This results in a 4-day aquisition time, if a total of 8 complex points in the ^{13}C (t1, t3) dimensions and 64 complex points in the indirect ^{1}H (t2) dimension are collected. As in the 3D ^{13}C HCCH-TOCSY, aliasing of frequencies in both ^{13}C dimensions is required, as well as linear prediction of the ^{13}C time-domain signals, to obtain sufficient digital resolution to make the four-dimensional approach viable. The 8 complex points are extended to a total of 12, using the linear prediction function of the NMR processing program FELIX (Biosym), or to 16 points using a 'mirror image' linear prediction method (Zhu & Bax, 1990). The final processed data comprise a 4D matrix of size 256 x 128 x 32 x 32 real points. The ^{13}C dimensions are designed to produce a change of sign for signals which are aliased an odd number of spectral widths away from their actual chemical shift values. This information can be useful for eliminating ambiguity in assignment of crosspeaks.

The 4D $^{13}C/^{13}C$-edited NOESY data are most easily visualized as various combinations of 2D planes. This was accomplished through the use of macros developed 'in-house' for use with FELIX by Andrew Lee. A representative set of slices from the 4D $^{13}C/^{13}C$-edited NOESY spectrum of the NTRC receiver domain is shown in Figure 1.12(B-D). The boxed crosspeak is seen from three 2D perspectives: in a ^{13}C-^{1}H plane (B), a ^{13}C-^{13}C plane (C), and a ^{1}H-^{1}H plane (D). The slice of panel B is defined by a pair of chemical shift values, which correspond to the ^{13}C and ^{1}H chemical shifts of the large peak indicated with an arrow. This corresponds to a diagonal peak in the 4D spectrum (i.e., D1=D2 and D3=D4). All other peaks in that plane represent NOE correlations to the proton which has the ^{13}C and ^{1}H chemical shifts of the diagonal peak. One of these,

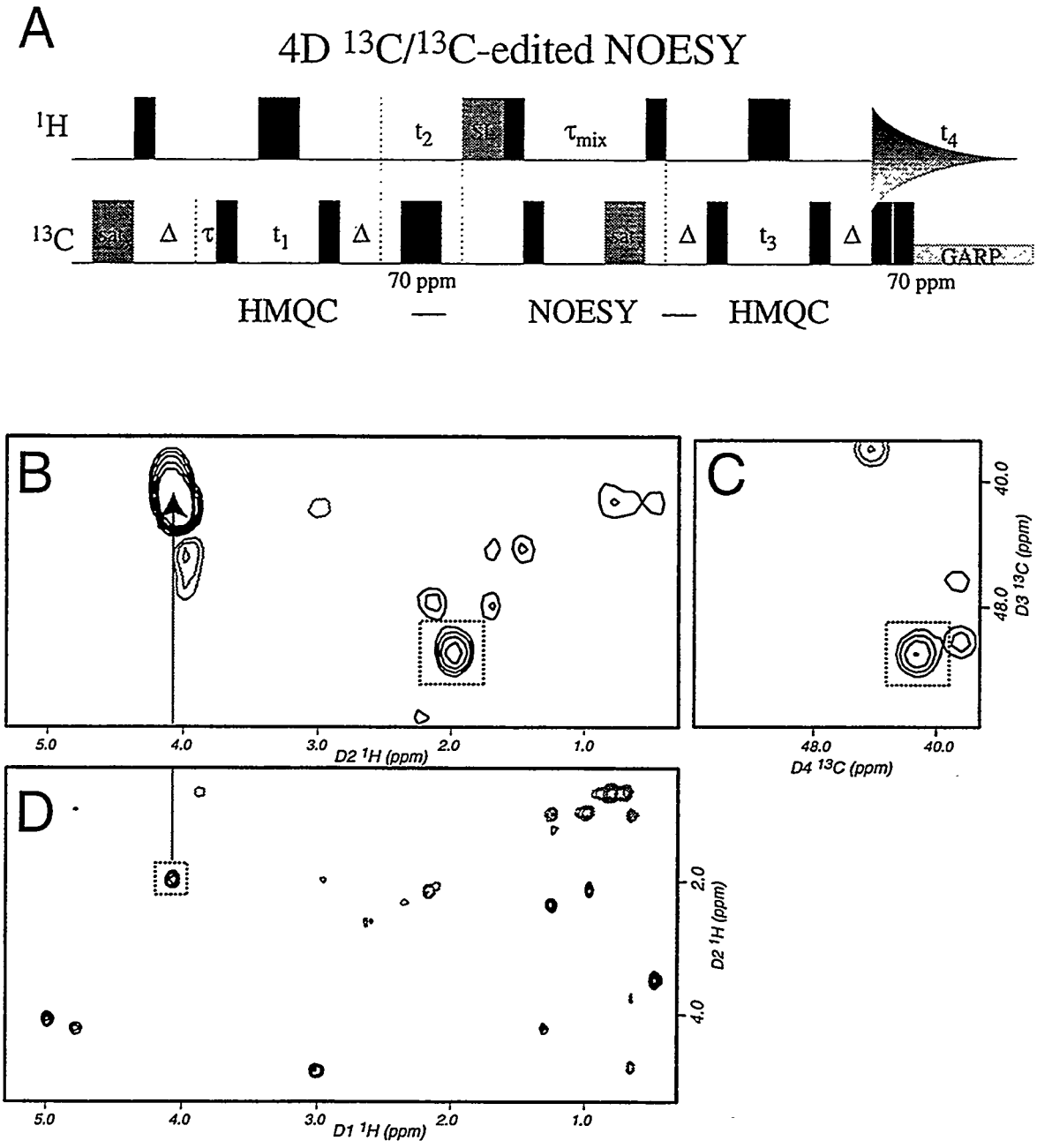


Figure 1.12: A. Diagram for 4D HCCH-NOESY. Spectrum is recorded in 4 days, collecting 8 complex points in each ^{13}C (t_1 and t_3) dimension, and 64 complex points in the t_2 (^1H) dimension. A pattern of aliasing similar to that of the HCCH-TOCSY is used in the ^{13}C dimensions, resulting in both positive (black) and negative (gray) peaks. B-D. Three different 2D slices from the 4D HCCH-NOESY spectrum of NTRC receiver domain, all of which contain the same (boxed) crosspeak. For a detailed description of relationship between the three slices and discussion of the analysis of the 4D HCCH-NOESY consult the text.

indicated with a box, is also viewed as a ^{1}H-^{1}H crosspeak in panel D. This 2D slice of the 4D is determined by the two ^{13}C chemical shifts corresponding to this crosspeak, which can be seen in panel C. This ^{13}C-^{13}C slice is likewise determined by the two ^{1}H chemical shifts of this crosspeak.

The initial approach to analyzing the 4D $^{13}C/^{13}C$-edited NOESY involves the systematic manual peakpicking of all crosspeaks in the spectrum to produce a database of chemical shift values, crosspeak intensities, and crosspeak sign. Starting from a ^{1}H-^{1}H slice and picking a crosspeak with a mouse-driven cursor activates a FELIX user macro to read in the ^{13}C-^{13}C slice which has been selected by the ^{1}H-^{1}H values of the picked peak. The appropriate ^{13}C-^{13}C peak is then selected, and all chemical shift values are stored to the database, along with the sign and qualitative intensity of the crosspeak. The 4D $^{13}C/^{13}C$-edited NOESY spectrum of the 124-residue protein which is shown resulted in a database of more than 1600 individual crosspeaks. This labor-intensive process should soon be supplemented with software which will at least partially automate peakpicking of 3D and 4D data, but the current version of FELIX software does not contain this capability.

Generation of Distance Restraints

As indicated in the Introduction, the focus of NMR-based methods for structure determination is the generation of interproton distance information, based upon NOESY experiments. For peptides and small proteins, the 2D NOESY can provide enough unambiguous information to define a structure, and in these cases the generation of distance restraints is a simple process that directly follows the sequential assignment of resonances. For larger proteins which require heteronuclear NMR methods the generation of distance restraints can be more complex, especially when the four-dimensional NOESY experiments are used. Prior to a brief description of distance geometry techniques a discussion of restraint generation is given, although distance geometry structure calculations are an integral part of this process.

To determine the structure of the NTRC receiver domain (Chapter 4), a large number of restraints were required for defining both the individual elements of secondary structure as well as the three-dimensional arrangement of the α-helices and β-sheet. The 3D ^{15}N-separated NOESY provided many of the helical NOEs, as well as a great number of the cross-strand connectivities of the parallel β-sheet, as shown in Figures 4.4 and 4.6. However, the majority of long-range connectivities needed to determine the global fold of the protein were derived from the database of 1617 4D $^{13}C/^{13}C$-edited NOESY crosspeaks. An iterative procedure for comparing the 4D NOEs with the table of chemical shift assignments and screening potential restraints against the preliminary structures was employed.

The first step in this procedure is to systematically evaluate every 4D NOE in terms of the potential assignments for the two protons (usually referred to as 'starting' and 'destination'). This involves comparing each pair of ^{1}H and ^{13}C chemical shifts in the 4D NOE list with the list of all known assignments for the protein. This task was performed using the spreadsheet program Excel, v. 4.0 (Microsoft) on a Macintosh Centris 650 personal computer (Apple Computer). A macro originally written by Andrew Lee in the Excel macro language was modified to work with the data structure of the assignment table for the NTRC receiver domain. The extract3.mac Excel macro can be found in Appendix A. The computation of all matches between the 1617 4D NOEs and the table of assignments required approximately 2 hr. to complete. The resulting Excel worksheet, extract3.wsh, was a 1.7 Mbyte file. A hard copy of this list of possible assignments was visually inspected to obtain a set of restraints with which to start structure calculations

Initially, a small number (<100) of long-range restraints derived from 4D NOEs were used; this is due to the degeneracy still present in the pairs of ^{1}H and ^{13}C chemical shifts defining each proton involved in an NOE. If two protons have the same ^{1}H and ^{13}C chemical shifts, any NOE involving either of the two protons cannot be unambiguously assigned without some structural knowledge of the geometry of the two possible

interactions. Once a low-resolution structure of the molecule is available, a number of ambiguous assignments can be resolved, if all but one of the combinations of proton pairs are clearly out of the range of NOE transfer (>> 5 Å).

To systematically screen potential restraints, the list of possible 4D NOE assignments contained in the extract3.wsh file was exported to a Unix workstation and used as input for a simple calculation by the program X-PLOR (Brünger, 1992). This routine takes the list of possible assignments, which have been reformatted so that the X-PLOR program can read it as a list of pairs of atoms, and calculates distances for each possible NOE from a protein structure coordinate file. In this way, the structures calculated at one stage of refinement are used to identify more restraints to be included in the subsequent stages of refinement.

Distance Geometry Methods

Two different programs were used for the calculation of structures from NOE restraints with the *metric matrix distance geometry* method. The DGII program, part of the NMRchitect module of the InsightII package (Biosym), was used in calculations of the peptide ApaBPTI and is described in Chapter 2. For the determination of the structure of the NTRC receiver domain described in Chapter 4, the program X-PLOR (Brünger, 1992) was used, and the approach taken for those calculations will be described here. X-PLOR is, at this time, the most commonly used program for calculation of solution structures of proteins from NMR data and refinement of biomolecular structures from x-ray diffraction data.

Metric matrix distance geometry (or just distance geometry) relies on the ability of a large number of interatom distances to define a set of atomic coordinates which are largely consistent with the distances. The set of NOE-derived distances is used to construct an (N x N) *bounds matrix*, where N is the number of atoms in the system, and the off-diagonal elements of the matrix represent upper- and lower-bounds of the corresponding interatomic

distances. The experimentally derived distances and covalently defined distance ranges are used to tighten the remaining elements of the table through the application of the triangle inequality, a process called bounds smoothing. Next a complete set of interatomic distances is randomly chosen within the ranges specified by the bounds matrix. These distances are then projected onto a three-dimensional space, resulting in a set of atomic coordinates, in a process known as embedding. A complete mathematical justification for the the embedding procedure has been described previously (Havel et al., 1983). These coordinates are then regularized and refined against the experimental restraints and a simple repulsive van der Waals term to give a final structure.

The standard protocol used (Nilges et al., 1988) consists of four steps, *dg_sub_embed*, *dgsa*, *refine*, and *accept,* each contained in a separate set of X-PLOR calculations requiring its own input file. The input files contain all of the commands and parameters to control the X-PLOR program and direct the flow and format of output generated. Output from each step consists of a log file (e.g. *dgsa.out*), as well as a set of coordinate files in Brookhaven Protein Data Bank (PDB) format. Any number of structures can be calculated, with computation time being the primary determinant. In the intermediate stages of refinement 10 structures were typically calculated, requiring approximately 15-20 hr. on a Silicon Graphics Indigo workstation with no competing processes.

The *dg_sub_embed* routine performs a substructure embed of the protein based on the smoothed bounds matrix. The method of substructure embedding (Havel & Wüthrich, 1984) was developed to allow the more computationally efficient generation of starting structures by including only a subset of the atoms of each residue in the actual embedding procedure, and then fitting the remaining atoms based on a template structure. The output of *dg_sub_embed* is a set of coordinates for the protein which are generally consistent with the selected distances, but the corresponding structure will have poor covalent geometry.

The *dgsa* routine takes the coordinate files generated by *dg_sub_embed* along with the restraint file as input and uses high temperature molecular dynamics to allow significant rearrangement of atoms to attain better agreement with the restraints and covalent geometry. After a period of dynamics at high temperature, the molecule is allowed to 'cool' or anneal, producing a conformation near the global minimum of the potential function. Another set of PDB coordinates is written at the end of the *dgsa* calculation.

The *refine* routine performs another simulated annealing on the *dgsa* structures followed by a short Powell minimization, only to fix relatively small violations of restraints or covalent terms, rather than producing large-scale rearrangements. This is the final modification to the coordinates. Following the *refine* routine is the *accept* routine, which simply evaluates the final structures in terms of restraint violations and values of the terms of the error function. It applies a user-defined cutoff for acceptable structures, and generates another set of coordinate files for the structures which pass the acceptance criteria. In the case of the NTRC receiver domain, a total of 30 X-PLOR structures were calculated in the final (35th) round of refinement, and 20 of them were accepted and analyzed as an ensemble of structures consistent with the 917 experimental restraints.

This approach to distance geometry calculation differs from others (like DGII) in the absence of full-atom embedding and metrization techniques (Havel & Wüthrich, 1984) in the interest of computational efficiency. Other approaches to structure determination are commonly used as well, usually relying on a pure simulated annealing approach from random starting structures, in contrast to the hybrid distance geometry/simulated annealing (DG/SA) approach described here or the DGII method which uses only low temperature dynamics in optimizing the embedded structures. Comparison of these various approaches has been useful both in analyzing the conformational sampling abilities of each relative to the others (Kuszewski et al., 1992), and also in demonstrating the general ability of NMR-based methods to produce high-quality representations of protein structures, establishing NMR as a strong partner to crystallographic methods in structural biology.

Chapter 2 Antipeptide Antibodies

Introduction

Since the introduction of techniques for production of monoclonal antibodies with hybridoma cells in 1975 (Kohler & Milstein, 1975), the impact on all areas of medicine and biochemistry has been enormous. By combining antibody-producing spleen cells from mice with murine myelomas, it is possible to produce immortal cell lines, or hybridomas, each of which produces a single antibody with a potentially unique binding specificity. The ability of antibody molecules to recognize large protein surfaces and small organic molecule antigens has been probed with structural studies (Davies et al., 1988; Haynes et al., 1994; Zhou et al., 1994a), and antibodies produced with small molecule antigens have been shown to catalyze a number of chemical reactions (Lerner et al., 1991). Because antibodies are exquisitely specific devices of molecular recognition, their development as potential therapeutics, engineered enzymes, or 'abzymes', as well as their application to questions of protein structure have been widespread.

By immunizing mice or rabbits with protein or peptide, many high affinity monoclonal and polyclonal antibodies have been produced, and it has been shown that antibodies produced from peptide fragments of a protein will cross-react with the native protein (Sutcliffe et al., 1980; Walter et al., 1980). In most cases, the immunogenic agent has been a small, unstructured polypeptide, often tethered to a 'carrier protein'. The carrier protein is needed to confer immunogenicity to a molecule which might otherwise be unable to incur a significant immune response. Given the unstructured nature of such small (<20 amino acids) peptides, the ability of antipeptide antibodies to recognize native protein must be based on primary structure and the constituent chemical functionality, rather than on the nature of the polypeptide fold in that region of the native protein (Dyson et al., 1988). If, however, a small but structured peptide antigen were used, it might be possible to generate antibodies which not only bind specifically to an amino acid sequence, but also be selective for its folding geometry, i.e. an α-helix or β-hairpin binding antibody.

To test this hypothesis, a small peptide antigen with a highly stable regular secondary structure was designed, based on the bee venom peptide apamin. Apamin is an 18 amino acid peptide with two disulfide bonds, stabilizing two turns of α-helix (Pease & Wemmer, 1988). It has been previously shown that amino acid substitutions in the helix may be made without disturbing the overall folded structure, provided the two disulfide bonds are left intact and in the same positions in the amino acid sequence (Pease et al., 1990). By replacing the helix in apamin with residues from a larger protein, for example BPTI, which are known to adopt a helical conformation in the native form of the protein, it was hoped that a small structured mimic of that region of the protein would be produced. Antibodies produced by immunization with this peptide antigen were selected for their ability to recognize the helix in both the peptide mimic and the native protein.

Peptide Design

The helix of apamin is contained in the C-terminal 10 residues of the peptide, and is stabilized by two disulfide bonds. The bovine pancratic trypsin inhibitor (BPTI) protein contains a C-terminal helix of similar size, also stabilized by two disulfide bonds. The structure of BPTI has been determined by x-ray crystallography to high resolution, and pure BPTI is availible commercially. Figure 2.1 shows the amino acid sequences of apamin, the C-terminal residues of BPTI and a hybrid sequence combining the C-terminal helix of BPTI with the first 8 residues of apamin. Conveniently, the cysteine residues of BPTI fit precisely into the arrangement required for proper formation of an apamin-like fold. By subsituting the residues of ApaBPTI into the three-dimensional structure of apamin using the Insight II program (Biosym), it could be seen that the solvent-accessible surface of the C-terminal helix in BPTI was also fully exposed in the model of the hybrid sequence peptide, Apa-BPTI.

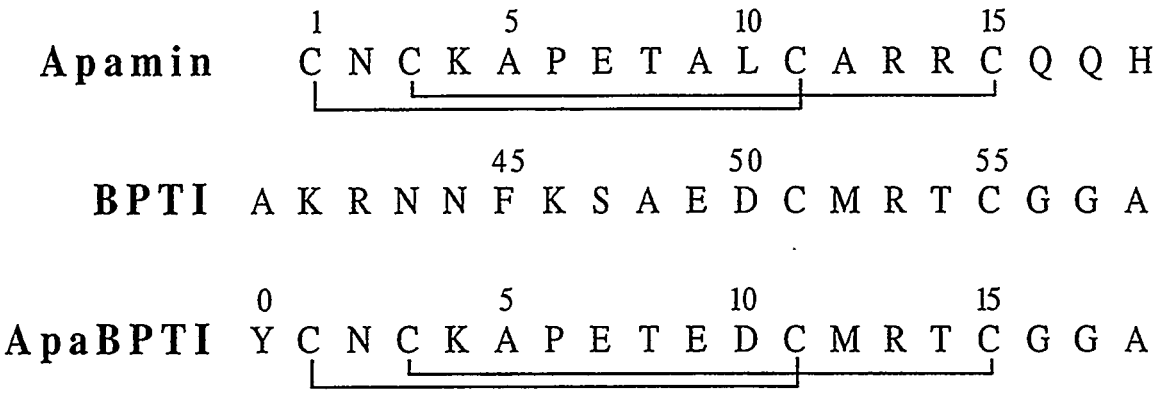


Figure 2.1: Amino acid sequences of apamin, bovine pancreatic trypsin inhibitor (BPTI), and ApaBPTI. Amino acid types are indicated with the one-letter abbreviation, residue numbers are indicated above each sequence. Disulfide bonds of apamin and ApaBPTI are indcated with horizontal lines connecting the corresponding cysteine residues. C55 and C51 of BPTI form disulfide bonds with cysteines residues in portions of the sequence not displayed here.

Two versions of ApaBPTI were synthesized, with and without an N-terminal tyrosine residue. The addition of a tyrosine allows quantitation of peptide concentration in solution by measurement of the UV absorbance at 276 nm. Although other apamin-derived hybrid-sequenced peptides have been studied (Pease et al., 1990), none included additional residues at the N-terminus, preceding the first cysteine residue. The two forms display nearly identical NMR spectra, and the addition of an N-terminal residue does not appear to inhibit folding. This result may prove useful, if similar apamin-derived peptides are included as N-terminus-linked subunits in larger constructs.

Peptide Synthesis and Purification

The peptide Apa-BPTI was synthesized on an Applied Biosystems ABI 430A automated peptide synthesizer, using standard Fmoc chemistry (Barany & Merrifield, 1980). Amino acid activation was accomplished by mixing 1mmole of 9-fluorenyl methoxycarbonyl (Fmoc)-protected amino acid with 1mL of 1M hydroxybenzotriazole (HOBt) in N-methyl pyrrollidinone (NMP) and 1mL 1M dicyclohexylcarbodiimide (DCC) for 30-60 min. This reaction produces a white precipitate of dicyclohexylurea (DCU) in the activation vessel, a helpful sign that the amino acid is successfully activated and ready to be added to the peptide-resin. While the activation reaction proceeds, piperidine is added to the peptide-resin in the reaction vessel to remove the Fmoc protecting group from the amino (N) terminus of the previously added residue. HOBt-activated amino acid is transferred to the reaction vessel and coupled to the deprotected N-terminus for two hours with constant vortexing.

After synthesis is complete the peptide resin is washed extensively in methanol to remove remaining NMP and piperidine, and allowed to dry completely before cleavage/deprotection. A general protocol for removal of the peptide from the resin and deprotection of amino acid sidechains was obtained from Dr. David King in the Department of Molecular and Cell Biology at U. C. Berkeley. Peptide resin (100-150 mg) is placed in

2 mL of Reagent K (King et al., 1990), a cocktail containing TFA, 5% phenol, 1.25% water, 2.5% thioanisole and 2.5% dithioethane. The mixture is incubated with regular stirring for 4-6 hours, after which the resin is vacuum filtered with a glass frit funnel. The resin is thoroughly washed with TFA, and the filtrate is reduced in volume to approximately 1 mL using a rotoevaporator. The highly concentrated TFA-peptide solution is forcefully squirted from a pasteur pipette into a conical glass centrifuge tube containing chilled methyl-*t*-butyl ether. This precipitates the peptide while keeping many of the protecting groups in solution. The resulting fluffy white precipitate is spun down in a tabletop centrifuge and the ether is decanted. The peptide is resuspended in fresh ether and spun down again, up to an additional six times to completely remove the TFA and other components of the cocktail, as well as protecting groups. The peptide pellet is air dried completely and pulverized into a fine powder before redissolving in water for HPLC purification. Peptide is purified by reversed phase HPLC, using an acetonitrile gradient in water containing 0.1% TFA as mobile phase and C18 stationary phase.

Apa-BPTI was synthesized using 0.25 mmole of Fmoc-Ala-Wang resin (Novabiochem) as starting material. A total of 485 mg of crude peptide was recovered, which was 95 % of the theoretical yield. Crude peptide (100 mg) was dissolved in 100 mL 20 mM Tris buffer, pH 8.0, with 10 mL β-mercaptoethanol, and incubated at room temperature for 1 week, to allow oxidation of disulfide bonds. The solution was concentrated using Amicon ultrafiltration with a YM1 membrane (500 Da cutoff). Oxidized peptide was HPLC purified on an IBM LC/9533 HPLC using a Waters preparative μbondapak C18 column, or on a Hewlett Packard Series 1050 HPLC using a Vydac semipreparative C18 peptide and protein column. Peptide was eluted using a gradient of 6% to 18% acetonitrile over 30 minutes, at flowrates of 15 mL/min. or 2 mL/min. HPLC fractions were dried in a Speed Vac concentrator (Savant) and stored at -20°C. NMR samples were prepared by dissolving dry peptide in either 90% H_2O/ 10% D_2O or 100% D_2O. Sample pH was adjusted to 2.75 with addition of 1M HCl or NaOH.

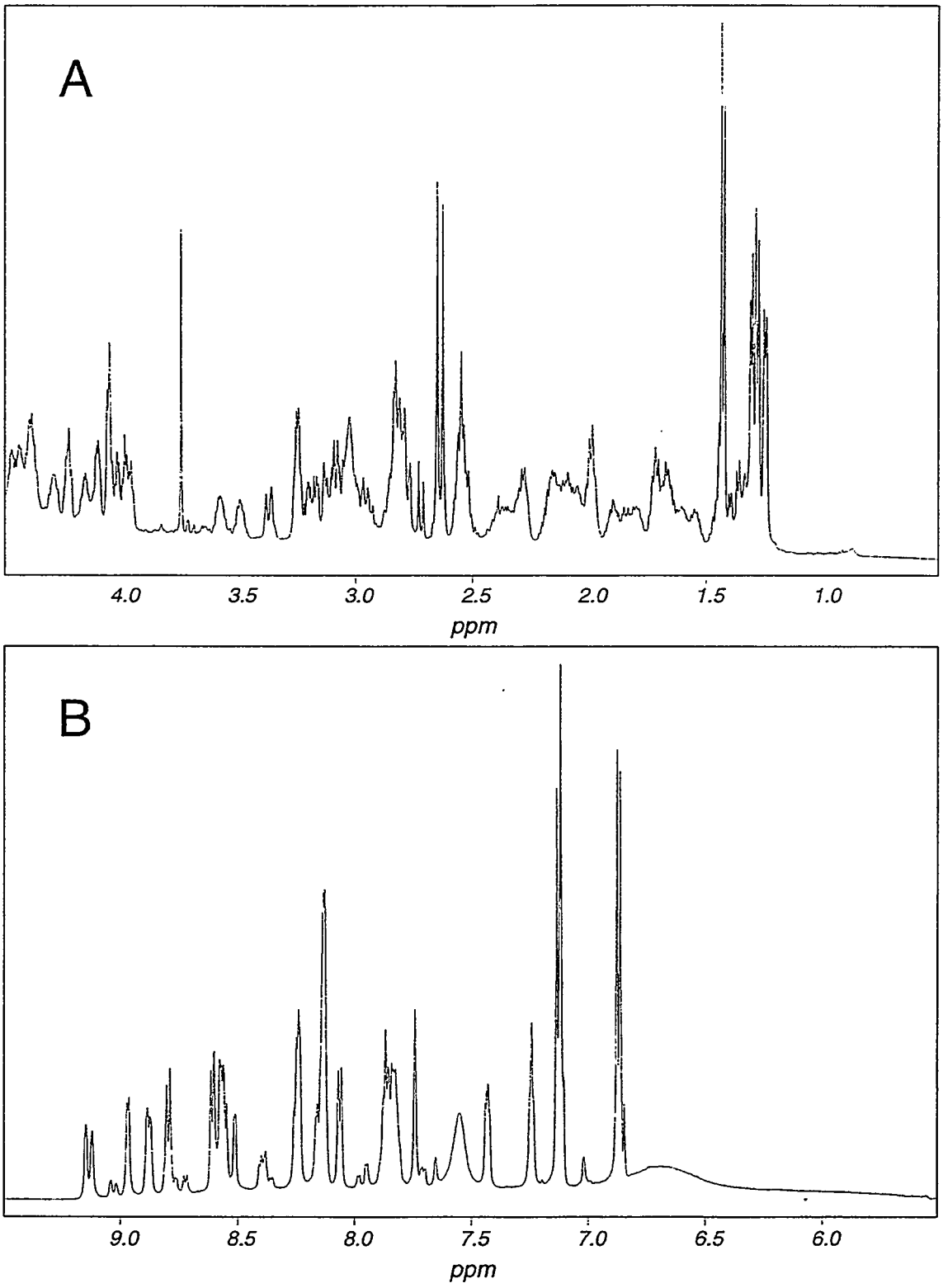


Figure 2.2: 600 MHz ^{1}H spectrum of ApaBPTI, 10 mM, collected at 10° C, pH 2.75. Panel A shows the upfield half of the spectrum, including all aliphatic resonances, and panel B shows the amide and aromatic protons of ApaBPTI.

Peptide Structure Determination

To confirm that Apa-BPTI provides a good mimic for the C-terminal helix of BPTI, 2D ^1H NMR and distance geometry were used to determine the three-dimensional structure of the peptide. First, 2D TOCSY and NOESY spectra were used to assign all ^1H resonances in the peptide; then the NOESY spectrum was completely assigned, identifying proton pairs with less than a 5 Å separation by the presence of a crosspeak at their respective frequencies. This distance information and the covalent structure of the peptide were used as input for a distance geometry algorithm, which produced a structure for the peptide which is consistent with the NMR data.

Resonance Assignments: A proton NMR spectrum of ApaBPTI is shown in figure 2.2. NMR data were acquired on a Bruker AMX-600 NMR spectrometer at 10° C. A 10mM sample of Apa-BPTI, pH 2.75, was prepared as described and 2D NOESY (Anil-Kumar et al., 1980) and TOCSY (Bax & Davis, 1985) experiments were collected. The spectral width in all experiments was 6024 Hz, and 2048 total data points were collected in t2. 64 scans were collected for each t1 increment. Solvent suppression was achieved by low power presaturation of the water resonance for 2 s. before each scan. Quadrature in the t1 dimension was achieved by time-proportional phase incrementation (TPPI) (Drobny et al., 1979). For the NOESY experiment in H_2O, a mixing period of 250 ms was used and 700 t1 points were taken. The TOCSY experiment utilized an MLEV-17y pulse train for the spin-lock for isotropic mixing, with a duration of 90 ms. NMR data were processed with Felix 2.14β (Hare Research) on Silicon Graphics workstations. Time domain signals were apodized with a skewed sinebell function, with a phase shift of 75° and skew value of 0.5. Individual rows were baseline corrected with a fifth-order polynomial fit to preselected baseline points.

Amide exchange information was obtained by examining 1D spectra of peptide at least 2 hr. after being dissolved in D_2O. The well-dispersed NH region (see figure 2.2)

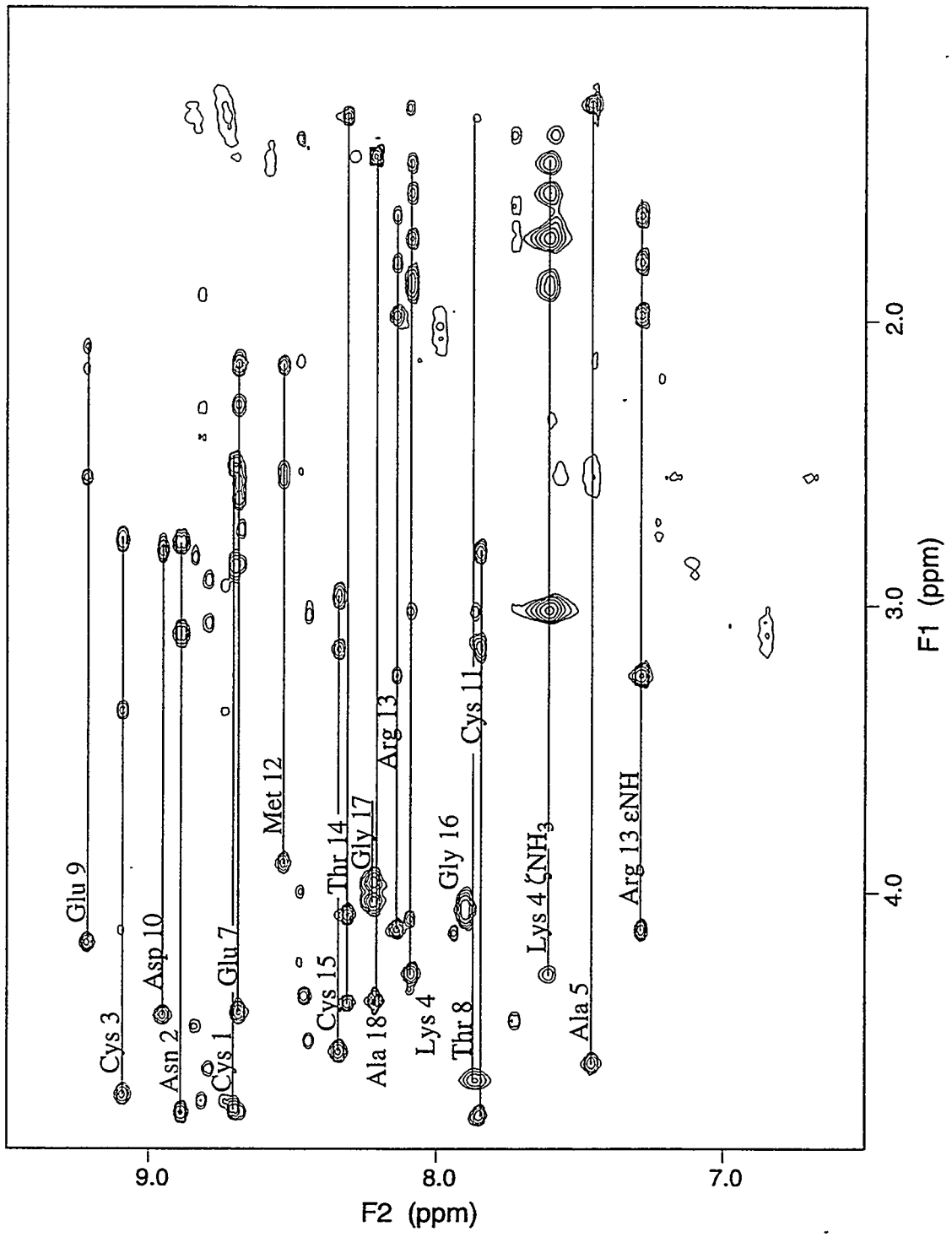


Figure 2.3: Amide to αH and sidechain region of the HOHAHA spectrum of ApaBPTI in H_2O at 10° C, pH 2.75. Residue assignments are indicated on the horizontal lines connecting crosspeaks of individual spin systems.

provided unambiguous identification of slowly exchanging amide protons. Amide protons which were visible at least 6 hr. after dissolving in D_2O were assumed to be in hydrogen bonds if the pattern of NOEs suggested an unambiguous acceptor.

The standard method used to completely assign the proton resonances of ApaBPTI relies on sequential NOE information connecting adjacent amino acid residues, primarily through the αH and NH protons (Wüthrich, 1986). By using the TOCSY experiment to link backbone amide resonances with the corresponding sidechain protons, complete identification of spin system types is achieved; a region of the TOCSY spectrum of ApaBPTI is shown in Figure 2.3. The pattern of crosspeaks observed at each NH chemical shift allows classification of the spin system according to the number and chemical shifts of its sidechain resonances. The NOESY experiment is used to connect the spin systems identified in the TOCSY which are adjacent in the amino acid sequence. NH-NH, αH-NH, or βH-NH crosspeaks in the NOESY spectrum provide the sequential connections between spin systems, allowing complete sequence-specific assignment of every proton in the peptide. The resonance assignments of ApaBPTI can be found in Appendix B.

The NH-NH region of the NOESY spectrum of Apa-BPTI, shown in Figure 2.4, provides unambiguous sequential information for connecting the spin systems in the helical portion of the peptide. Figure 2.5 shows the upfield region of the same spectrum, in which a number of sequential NOEs between NH protons and αH or βH's of the preceding residue are observed. Some intermediate-range αH-NH(i, i+3) connectivities indicative of α-helix are also found in Figure 2.5. These sequential and intermediate-range NOEs are summarized in Figure 2.6. Also indicated are amide protons which do not exchange rapidly when dry peptide is dissolved in D_2O. This is an indication that these backbone amides are involved in hydrogen bonding and probably the formation of secondary structure.

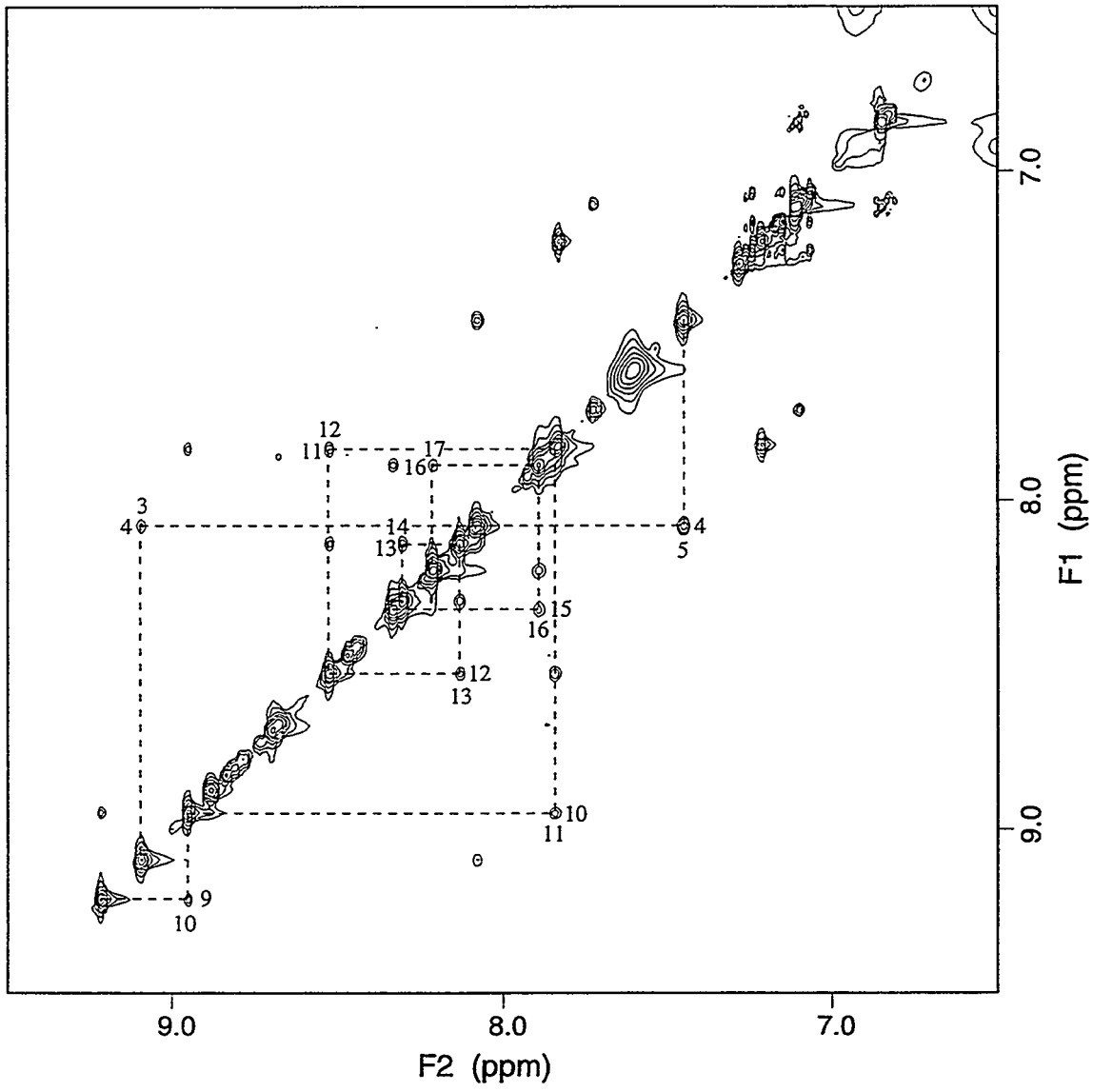


Figure 2.4: NH to NH region of the 250 ms NOESY spectrum of ApaBPTI in H_2O at 10° C, pH 2.75. Sequential connectivites are indicated with dashed lines, and crosspeaks are labeled with the adjacent residues numbers.

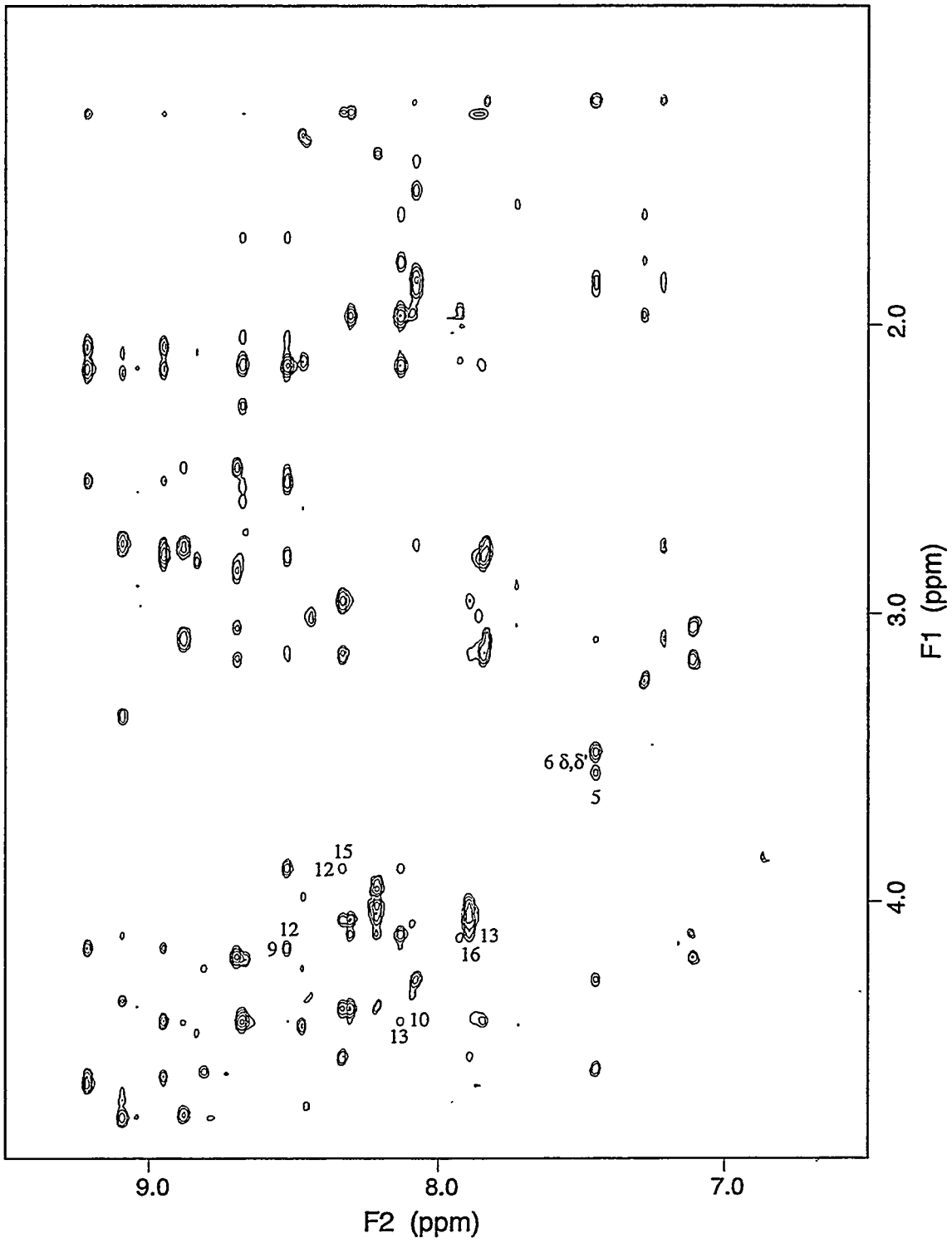


Figure 2.5: Amide to αH and sidechain region of the 250 ms NOESY spectrum of ApaBPTI in H_2O at 10° C, pH 2.75. Helical NOEs are indicated with the numbers of the corresponding residues.

Distance Geometry: The distance geometry calculation requires two types of input: the covalent structure of the molecule and the experimental restraints. The ApaBPTI molecule was created using the Builder module of Insight II, using the built-in library of amino acid structures. A molecular data file was created, which contained all of the covalent geometry information needed by the DGII program. Distance restraints were generated based upon crosspeaks from the NOESY spectrum which could be unambiguously assigned to a single pair of protons. Distance restraints were classified as either strong, medium or weak, based on the integrated volumes of the corresponding NOESY crosspeaks. All restraints were initially assigned lower bound distances of 1.8 Å and upper bounds of 3.0, 4.0 or 5.0 Å, corresponding to strong, medium and weak NOESY crosspeaks, respectively (Clore et al., 1986). The DGII program used these data to produce an ensemble of structures which are consistent with the experimental and covalent restraints; DGII is closely related to an earlier program, DISGEO (Havel & Wüthrich, 1984).

The process of structure calculation by distance geometry involves three basic steps: bounds smoothing, embedding, and optimization. First the distance contraints are loaded into the bounds matrix, which is an $n \times n$ matrix, where n is the number of atoms in the structure. This bounds matrix is then optimized by using the triangle inequality (and quadrangle inequality, if desired) to tighten upper or lower bounds for all atoms on the basis of the experimental data and covalent structure.

The embedding process randomly selects distances which fall between the upper and lower bounds for each atom pair and then projects these distances onto a three-dimensional structure. An optional procedure, called metrization, may be employed in the embedding step to better correlate the interatomic distances as they are being chosen. By resmoothing the bounds matrix after each individual distance is chosen, each set of embedded coordinates takes full advantage of the information contained in each random distance selection, to further tighten the limits of the distance space. After the projection

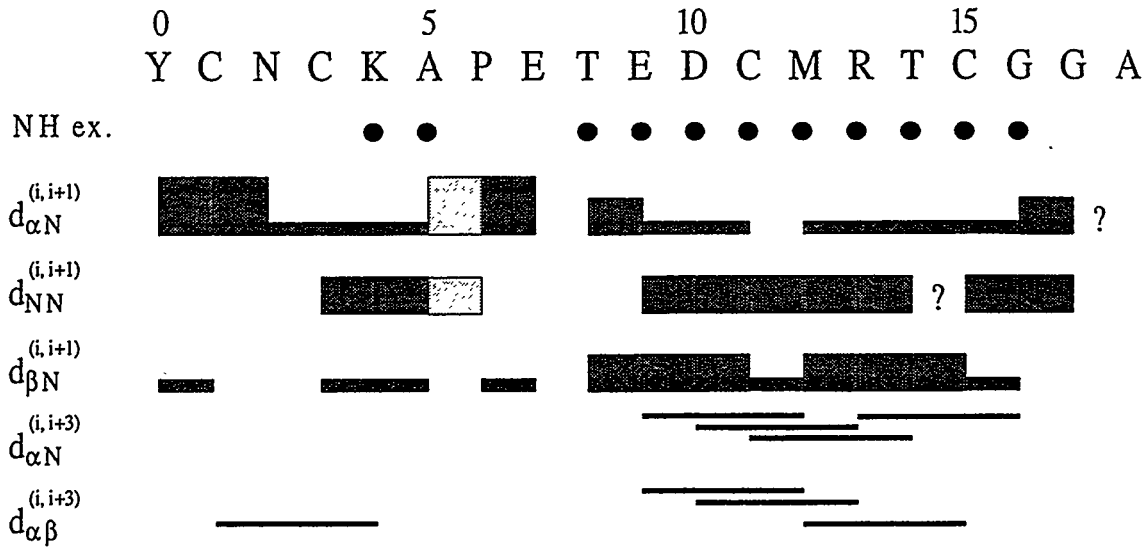


Figure 2.6: Sequential NOE diagram for Apa-BPTI. Shaded blocks indicate crosspeaks observed between protons of adjacent residues. Lightly shaded blocks indicate NOEs to Pro 6 δH's, in lieu of a backbone NH. Dots indicate backbone NH's which persist after redissolving in D_2O for more than 2 hr. Horizontal lines indicate NOEs observed between protons three residues apart in the sequence. Question marks indicate NOEs obscured by the diagonal or other peaks.

onto three-dimensional coordinates, the embedded molecule is 'majorized', using a least-squares fitting routine to obtain the closest agreement between the embedded coordinates and the trial distances.

After embedding, optimization of the molecule is needed to correct the improper covalent geometry and steric interations which are not considered in the generation of the starting molecule. The embedded coordinates are annealed and minimized against an error function which is dominated by the distance constraints and covalent structure. If the structure converges to an acceptable error level, it may then be minimized against a normal molecular force field to eliminate minor overlap or torsional anomalies. The embed and optimization steps are generally performed multiple times to generate an ensemble of structures displaying close agreement, i.e. low root mean square deviations (RMSDs) of atomic coordinates. Such a family of DG structures is evaluated to determine which regions of the structure are more or less well-defined; often a region of a protein which is highly flexible will display far fewer NOE's and subsequently appear as an undefined region in the family of DG structures. The lack of NOEs cannot be taken as conclusive evidence for local mobility, but dynamic behavior can be probed with studies of nuclear relaxation times of individual residues (Barbato et al., 1992).

A total of 135 restraints, including 123 NOE-derived distance restraints, 10 hydrogen bond restraints from D_2O-exchange data, and 2 disulfide dihedral restraints were used in calculating 10 DGII structures. A restraint file in DGII format is included in Appendix B. Triangle and tetrangle inequality metrization were performed on the initial bounds matrix, and prospective metrization was employed in the embedding process. Optimization of embeds was accomplished with simulated annealing at 300K for 20000 steps of 0.2 psec each, followed by up to 250 steps of conjugate gradients minimization, until a gradient of less than 0.01 was obtained. Of the 10 embeds, 7 with low energies and no restraint violations were chosen to form a family of convergent structures. A comparison of the 7 best Apa-BPTI structures, shown in Figure 2.7, reveals an average

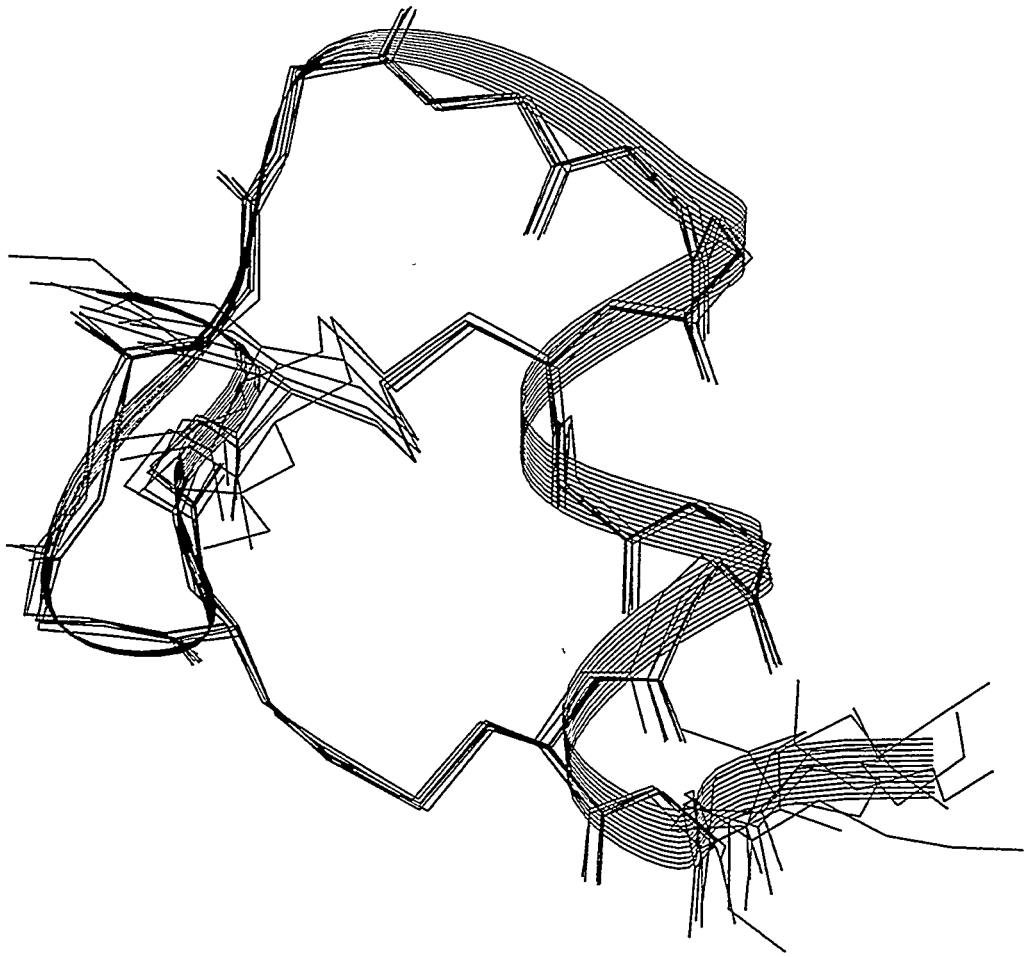


Figure 2.7: Family of 7 best structures out of 10 embeds from DGII. Individual structures were superimposed on the backbone atoms of residues 1-16 of the structure with the lowest RMSD to the average structure. The atoms used in the superimposition have an mean RMSD of 0.28 ± 0.22 Å for the 7 structures.

backbone RMSD of 0.28 Å for residues 1-16, while the lack of NOEs at the C-terminus produces no similarity between individual structures for residues 17 and 18.

Structural Comparison of ApaBPTI and BPTI

Figure 2.8 shows Molscript (Kraulis, 1991) ribbon diagrams of ApaBPTI (bottom) and BPTI (top). The coordinates of Brookhaven Protein Data Bank (PDB) file 5PTI were used to generate the diagram of BPTI. The structures are oriented similarly with respect to the C-terminal helix. It can be seen that the residues of the C-terminal helix of BPTI should be largely solvent exposed, with the sidechains of the cysteine residues dictating the region of primary interaction with the core of the protein. With this in mind, it would appear that the solvent-accessible surfaces of the C-terminal helix of BPTI and the ApaBPTI helix should be comprised of very similar groups.

Figure 2.9 shows two views of the family of ApaBPTI structures superimposed on the C-terminal helix of BPTI. The average RMSD of residues 10-17 of the ApaBPTI helix from the corresponding residues of BPTI is 1.00 ± 0.07 Å for the backbone atoms, and 2.01 ± 0.18 Å for all heavy atoms of those residues. The top and bottom views are related by a 90° rotation about the horizontal axis, and both indicate clearly the close agreement in positioning of the sidechains of the helices of BPTI and ApaBPTI. From the bottom view especially, it is apparent that the scaffold residues of ApaBPTI (0-8) should have no significant interactions with the sidechains of the helix. Based on these comparisons, the ApaBPTI peptide should present a very similar antigenic surface to the corresponding region of the BPTI structure.

Immunization and Antibody Production

Many approaches to antipeptide antibody production have been reported in the literature (Doolittle, 1986). Typically, small peptide fragments have relatively low immunogenicity, and additional steps must be taken to increase the animal's sensitivity to

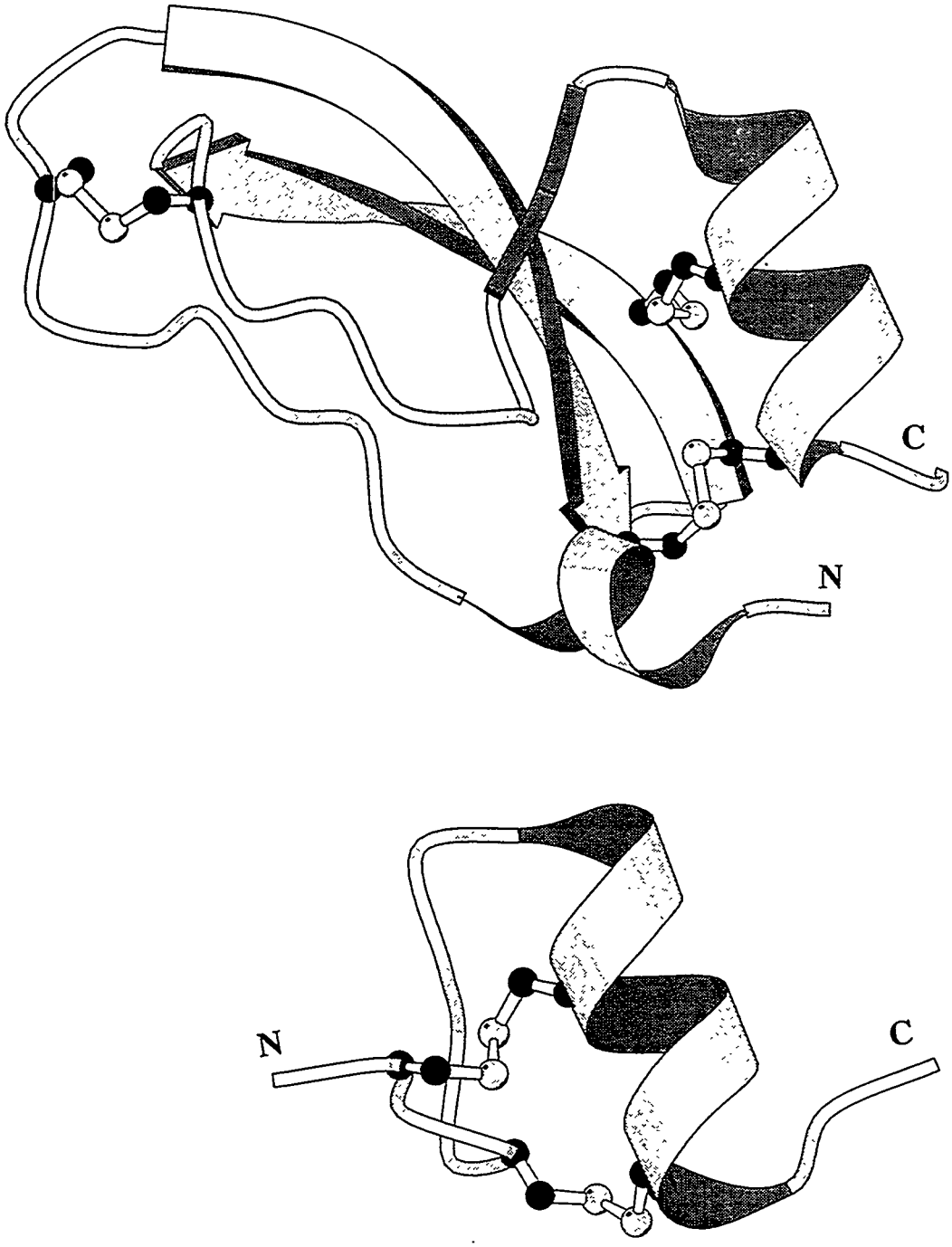


Figure 2.8: Ribbon diagrams of ApaBPTI and BPTI. The first structure of the ensemble of ApaBPTI DGII structures and the Brookhaven PDB file 5PTI were used to generated the diagrams. The molecules are oriented similarly with respect to the carboxy-terminal helix of BPTI and the helix of ApaBPTI.

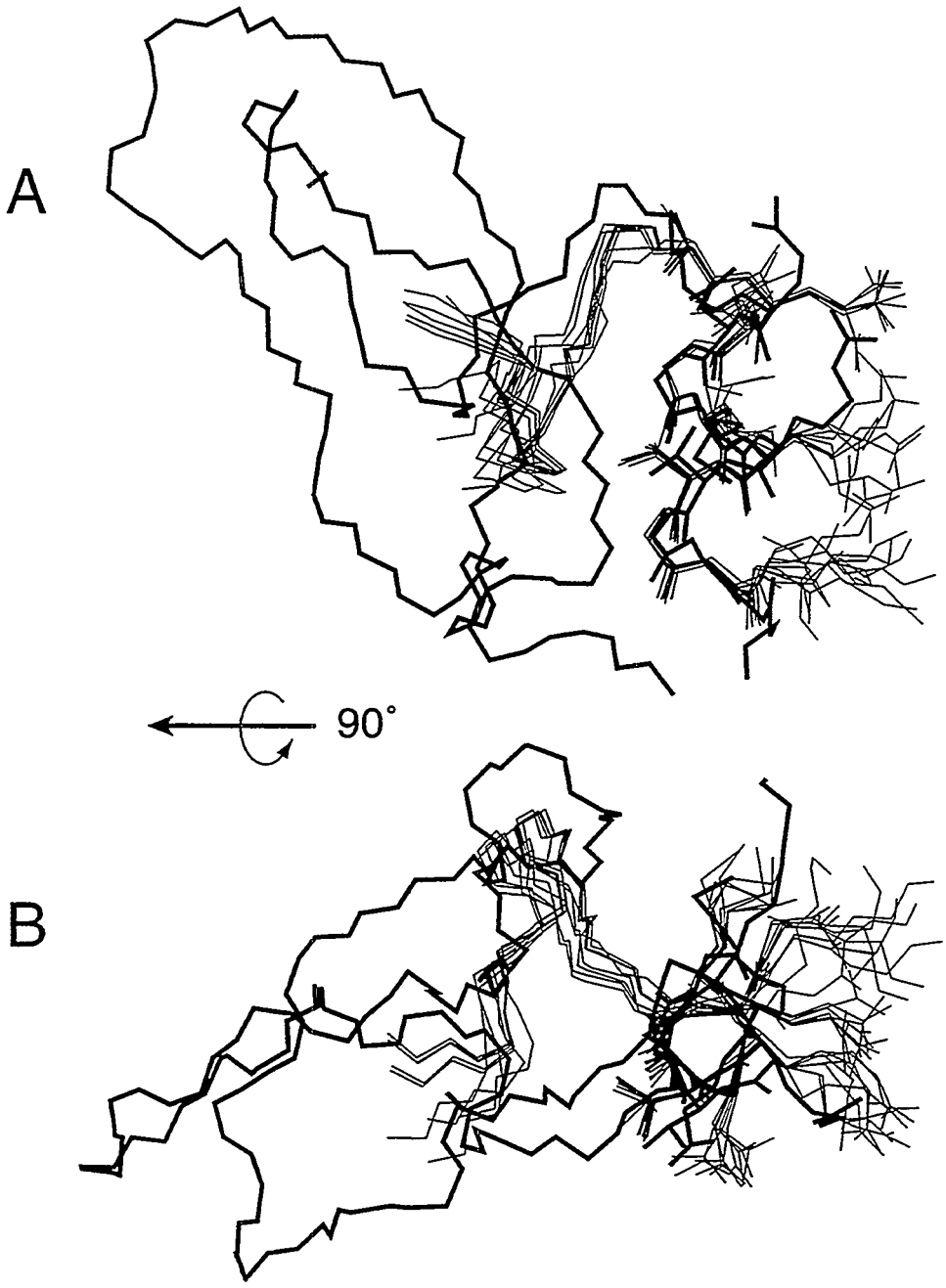


Figure 2.9: Direct comparison of family of ApaBPTI structures with BPTI. In panel A, the backbone traces of BPTI (heavy line) and the family of 7 ApaBPTI structures are superimposed using just residues of the helices: ApaBPTI: 9-16, BPTI: 49-56. The orientation is similar to that of Figure 2.8. In panel B, the view of panel A is rotated 90° about the horizontal axis to afford a view down the helix axis, from amino-terminal to carboxy-terminal.

the antigen. Often this is accomplished by chemically linking the peptide to a large protein with well established immunogenicity. The carrier protein-peptide complex is used as the immunizing agent. The primary drawback is the large amount of background noise in the immune response due to antibodies which recognize the carrier protein and not the small peptide. Because of the low abundance of antibodies with the desired functionality, extensive screening is required. It was hoped that the structured nature of the small Apa-BPTI peptide would render it a strong immunogen, eliminating the need for a carrier protein in immunization.

A standard protocol for murine antibody generation was followed (Ausubel et al., 1987). Apa-BPTI which had been fully oxidized, HPLC purified and structurally characterized was dissolved in phosphate buffered saline (PBS) solution at a concentration of 0.5 mg/mL. For the initial immunization of a mouse, the peptide solution was mixed with an equal volume of Freund's complete adjuvant. The mixture was repeatedly forced through an 18 ga. sterile needle to form a heavy suspension. Mice were immunized intraperitoneally (I.P.) with 50 - 150 μL of this suspension. For all subsequent (booster) immunizations, Freund's incomplete adjuvant was substituted for Freund's complete adjuvant, and the procedure was repeated similarly in all other respects.

Immunizations were spaced at least two weeks apart, and serum samples were collected from each mouse 3 days after each booster injection. Sera were collected from tail-bled mice which had been incubated in a light box for 10 min. After coagulation of the whole blood, the clot was removed with the wood end of a sterile cotton swab. The remaining serum was chilled on ice for 15 min. and then spun gently in a small tabletop centrifuge for 20 sec. The clear serum was pipetted from the top of the pelleted cells and frozen at -20C until needed for ELISA assays.

Antibody Binding Studies

To assess the level and specificity of antibodies produced, enzyme-linked immunosorbent assays (ELISA) were performed on sera collected from mice which had been immunized two or more times. 96-well microtiter plates were coated with 1-2μg of Apa-BPTI which had been covalently linked by its amino terminus or lysine sidechain to bovine serum albumin (BSA) with an MBS linker (Pierce). The sera were serially diluted in Tris-glycine buffer, pH 7, with 5% calf serum to prevent non-specific binding, and mixed with a competitive inhibitor, if used, and incubated for 15 min. at RT before loading onto the ELISA plate and incubated for 1 hr., RT. The plate was then washed with an automatic platewasher and 100 μL of a 1μg/mL anti-mouse γ-chain-specific antibody-alkaline phosphatase conjugate was applied to each well. After incubation at room temperature for 1 hour, the plate was again washed, and 100 μL of 1mg/mL nitrophenylphosphate (NPP) solution was applied to each well. After 1 hour, an automatic plate reader recorded the absorbance at 405 nm in each well.

Figure 2.10 shows a graph of ELISA data for polyclonal sera of a representative mouse. The presence of of antibodies which bind to ApaBPTI is clearly indicated by the strong ELISA response, as well as the high level of inhibition by ApaBPTI as a competitor. In addition, a competition assay with BPTI appears to lessen the response, suggesting the presence of some antibodies which recognize native BPTI. A fusion of spleens from this and another mouse was not successful in producing viable hybridomas.

Discussion

Although monoclonal antibodies with demonstrated affinity for BPTI were not obtained, the results observed with polyclonal sera indicated the presence of antibodies which can recognize the carboxy-terminal helix of BPTI. This implies the possibility of production of antibodies to structural epitopes, comprised of non-sequential residues. This may provide an improvement in our ability to manipulate immune responses to specific

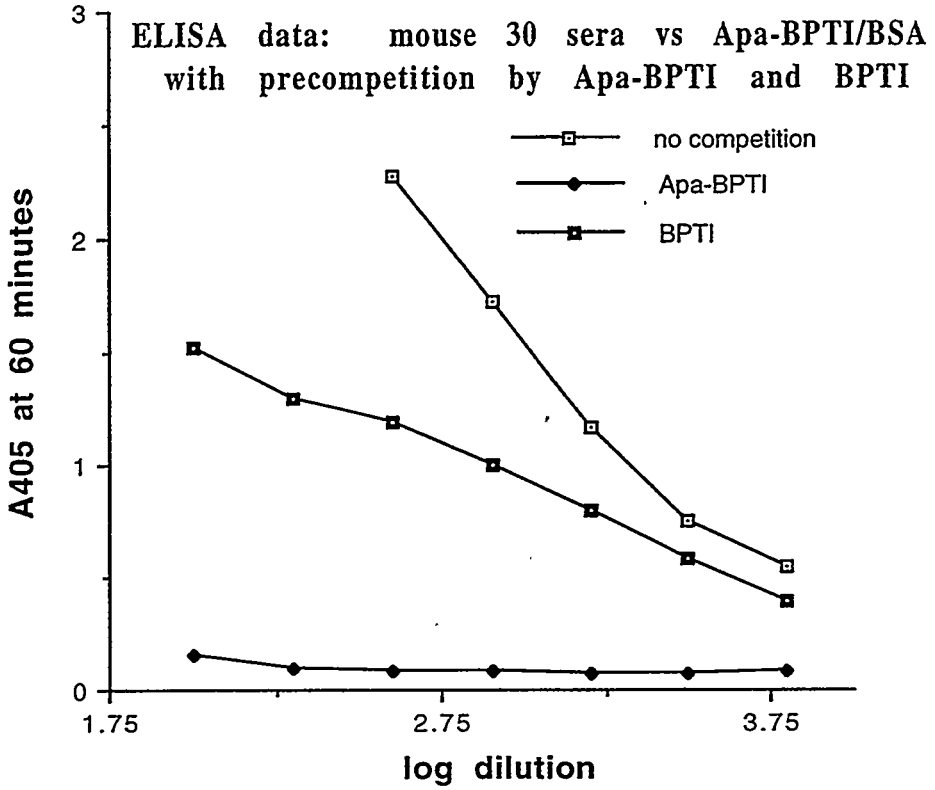


Figure 2.10: ELISA response to BSA-linked Apa-BPTI of polyclonal serum from mouse 30 after 3 immunizations with ApaBPTI. Serum was incubated with the competing molecule for 15 min. followed immediately by incubation on the ELISA plate for 1 hr. At the two lowest dilutions, the wells with no competitors had readings which were off the scale of the microtitre plate reader.

classes of antigens which share structural similarities but have highly variable sequences, like retroviral proteins.

Despite the indication of the presence of antibodies which cross-react with BPTI, assays of monoclonal antibodies produced by fusion of spleen cells with myelomas revealed no anti-ApaBPTI antibodies with high affinity for BPTI. An extensive study of the antigenic properties of apamin may provide a possible explanation for the lack of strong cross-reactivity (Defendini et al., 1990).

Defendini et al. made monoclonal antibodies to apamin using both the free peptide and a BSA conjugate as immunogens. By evaluating the cross-reactivity of the anti-apamin monoclonals with alanine-substituted apamin anaolgs, they were able to define the epitope of the apamin structure recognized by each antibody. Positions probed by alanine substitution include residues 2, 4, 7, 8, 10, 13, 14, 16, 17, and 18. A non-disulfide bonded apamin with acetomidomethyl protecting groups on the cysteine sidechains (4 Acm apamin) was also tested as an nonhelical control. The folded nature of the analog peptides was confirmed by CD, which showed a helical signal for all but the 4 Acm apamin peptide. By comparing levels of inhibition of apamin recognition by the alanine-substituted apamin analogs, the epitope recognized by each monoclonal was mapped.

A distinct difference was noted in the range of epitopes recognized by the antibodies elicited with free peptide and peptide conjugate. A total of 9 anti-apamin monoclonals and 11 anti-BSA-apamin conjugate monoclonals were analyzed for epitope specificity. Antibodies induced with a mixture of apamin and a BSA-apamin conjugate showed a range of epitopes which included all of the probed sequence positions, with the exception of residue 10 (analog L10A). These monoclonals were consistently most sensitive to substitution at position 7 (analog E7A). In contrast, all of the antibodies elicited with free apamin alone recognized primarily the residues of the top of the helix. Positions 10 and 13 were the primary epitopic determinants, with positions 7, 8, and 14 also contributing in almost all cases. These residues make up a contiguous solvent exposed surface on the

helix of apamin. In addition, while none of the anti-apamin (free peptide) antibodies recognized 4 Acm apamin, 2 of the 11 anti-BSA-apamin antibodies showed some ability to bind to the unfolded peptide.

While not all of the conclusions drawn by the authors of this study are convincing (e.g. speculation about T-cell vs. B-cell epitopes, antibody induced folding), two points have special interest. In agreement with the results obtained with ApaBPTI, apamin is shown to be immunogenic without being covalently linked to a carrier protein. Moreover, the lack of recognition of the unfolded 4 Acm apamin by all of the anti-apamin antibodies clearly indicates that the epitopes of apamin are comprised of surfaces of the helical peptide. Both of these results support the possibility of obtaining anti-ApaBPTI antibodies with the ability to bind native BPTI.

There is also a more unfortunate implication of the epitope mapping of apamin with respect to the anti-ApaBPTI study. The importance of residues 7 and 8 in the epitopes of apamin, if similar in ApaBPTI, might reduce the likelihood of cross-reactivity of anti-ApaBPTI antibodies with BPTI. Residues 7 and 8 of ApaBPTI are derived from the sequence of apamin, rather than BPTI (Figure 2.1), and the differences (E vs. S and T vs. A) may be significant enough to severely restrict the epitopes common to both structures.

Two approaches could be taken to improve the design of this experiment, based on the results of Defendini, et al. Redesign of the ApaBPTI peptide could either include the serine and alanine of BPTI at positions 7 an 8, respectively, or the helical residues of BPTI could be shifted down by three or four positions in the apamin construct, roughly one turn of helix. Another approach suggested by the apamin study would use a ApaBPTI-BSA conjugate for immunization to expand the possible epitopes to include all residues of the peptide. As long as a stable helical peptide results from a sequence including the additional residues of BPTI, immunization with this construct might present the most straightforward method for improving cross-reactivity without diluting the immune response with the use of a carrier protein conjugate.

Chapter 3 NMR Studies of the Ferric Uptake Regulation (Fur) Protein

Introduction

Iron is an essential nutrient for virtually all living systems, but its chemical reactivity stipulates that careful control be maintained over the cellular concentration of iron (Halliwell, 1988). In both prokaryotes and eukaryotes iron levels are regulated primarily by control of iron uptake (Neilands, 1990a). Because no systems for eliminating excess iron from cells are known, uptake regulation is thought to be the only way to control total cellular iron. Storage proteins like ferritin (in eukaryotes) are used to sequester excess iron and provide a secondary level of control. Transferrin proteins have been identified as primary iron transport agents in mammalian cells, but prokaryotic iron uptake occurs primarily through the use of siderophores (Neilands, 1981). Siderophores are small molecule metal chelators which are taken up by the cell when a specific receptor protein is activated.

Although all life forms exert uptake control over iron levels, a general distinction may be made between the mechanisms of controlling uptake in prokaryotes and eukaryotes. In mammalian systems, production of both the transferrin receptor and ferritin is controlled post-transcriptionally in an iron-dependent manner by the protein IREBP (iron responsive element binding protein). In low-iron conditions IREBP simultaneously represses translation of ferritin mRNA and enhances translation of the transferrin receptor by binding and stabilizing its mRNA at the iron responsive element (IRE) sequence (Klausner et al., 1993). In contrast, transcriptional control over the levels of many components of the iron uptake systems in bacteria is exerted by a single protein called Fur (ferric uptake regulator). The 147 residue Fur protein acts as an iron-dependent repressor protein, binding to the promoter sites of a number of (ferric) iron transport genes, including those for siderophore synthesis and transport, as well as to the *fur* promoter (Bagg & Neilands, 1987; de

Lorenzo et al., 1987; de Lorenzo et al., 1988a; de Lorenzo et al., 1988b). Fur also regulates transcription of a wide range of other genes including superoxide dismutases (*sodA* and *sodB*), succinate dehydrogenase, several bacterial toxins and the ferrous iron uptake system (Coy, 1993).

The implication that the Fur protein binds to both DNA and iron has been confirmed biochemically with experiments demonstrating that Fur acts as a classical repressor, blocking transcription in the presence of activating metal ions (Bagg & Neilands, 1987). The exact nature of either the DNA recognition or metal-binding properties of Fur have not been determined, and sequence analysis has not identified clear homology to known DNA-binding motifs. The Fur protein exists as a dimer at low concentration and appears to form tetramers at high concentrations (Coy, 1993). The functional domain structure of Fur has been probed biochemically using limited proteolysis (Coy & Neilands, 1991). The Fur protein is comprised of two domains, with the N-terminal half of the sequence containing the DNA-binding determinants, and the C-terminal domain containing the metal binding site. Previous NMR analysis has resulted in some partial resonance assignment information for the intact Fur protein, as well as a model for the structure and metal binding site (Saito & Williams, 1991; Saito et al., 1991). To develop a more accurate model for the structure and function of the Fur protein, we have continued to use NMR as a tool for developing structural information. We will also attempt to probe the binding of metal ions to the Fur protein and its metal binding domain with NMR.

Initial attempts to study the whole 147 amino acid Fur protein in this lab were largely unsuccessful due to the large ^1H linewidths of the 34 kDa dimer or 68 kDa tetramer. It was not possible to confirm the partial resonance assignments previously reported with the typical 2D NMR sequential assignment methods described in Chapters 1 and 2. To obtain accurate structural information about Fur, subsequent studies focused on the metal-binding domain of Fur, initially using samples prepared proteolytically from intact Fur protein. This domain, called FurT, is comprised of the C-terminal 71 amino acids of Fur

and has been shown to bind metal in a similar fashion to the full-length protein. In the later stages of NMR analysis, recombinant forms of this domain were overexpressed in bacteria and used to obtain higher levels of pure protein than could be obtained from tryptic digestion of Fur protein.

^{1}H NMR spectra of FurT show linewidths typical of a 16 kDa (or larger) protein, suggesting that the dimerization determinant of Fur lies in this region of the sequence, since the molecular weight of the FurT monomer is 8 kDa. Comparison of 1D and 2D ^{1}H and ^{15}N NMR spectra of full-length Fur and the FurT fragment provides a qualitative indication that the folded structure of the C-terminal domain is the same in both forms. It is hoped that elucidation of structural features of this domain will lead to some understanding of the manner in which the activity of Fur is controlled by iron binding. Examination of the amino acid sequence of Fur provides a few indications of likely metal binding residues. The Fur protein contains 12 histidines, and 9 of them are contained in the FurT fragment; however, mutational analysis at each of the histidines of Fur provides no clear indications of His-mediated iron binding (Coy, 1993), in agreement with NMR experiments. All 4 cysteine residues of Fur are found in the metal-binding domain. The only previously observed sequence motif identified in Fur is a -Cys-X-Y-Cys-Gly- found at residues 92-96. This sequence is part of the metal binding sites of several bacterial ferridoxins. Future experiments will address the possible role of these cysteines in the metal-binding activity of Fur.

This chapter presents the NMR analysis of the C-terminal metal-binding domain of the Fur protein, including nearly complete backbone ^{1}H and ^{15}N resonance assignments and a tentative secondary structure map of this domain. It is anticipated that this work will continue, resulting in complete ^{1}H, ^{15}N and ^{13}C assignments for the metal-binding domain as well as the complete secondary structure and the determination of the metal liganding residues.

Materials and Methods

Expression and Purification of Fur, Fur(77-147)(FurT), Fur(80-147)(FurT80) and Fur(88-147)(FurT88): The sequence of the Fur protein and the C-terminal domain constructs is shown in Table 3.1. The full-length Fur protein was produced as previously described (Wee et al., 1988). Preparation and purification of FurT from pure Fur protein has also been described (Coy & Neilands, 1991). The FurT80 and FurT88 constructs were cloned, expressed and purified in the Wemmer lab by Mark Coy as described below.

To construct an expression vector for FurT80 the oligonucleotide GATCGGCCATGGAACTGACACAGCAACATCAC, which contains, 5' to 3', a six basepair spacer, an NcoI site overlapping an initiation codon, and codons corresponding to residues 80-86 of the Fur sequence, and GATCGGCCATGGCTGGCCTGAAAGTAAC GCTT which has, reading 5' to 3', a five base spacer, a HinDIII restriction site, and seventeen bases complementary to the last part of the Fur structural gene, were synthesized. They were used to prepare a fragment which codes for a C-terminal fragment starting at amino acid 80 by amplification with the polymerase chain reaction (PCR) from the Fur expression vector pMON2064 (Wee et al., 1988) which had been linearized with HinDIII. The fragment was digested with an excess of HinDIII and NcoI and ligated into the expression vector pET28a+ (Novogen), which had been previously digested with HinDIII and NcoI and treated with calf alkaline phosphatase. The ligation mix was used to transform BL21(DE3), and the desired recombinants were selected by screening for the correct restriction digest map and by analysis of protein expression. A vector for expression of T88 was constructed in a similar manner, except that the oligonucleotide GATCGGCCA TGGATCACCTGATCTGCCTCGAC, which contains a six base spacer, an NcoI restriction site overlapping an initiation codon, and a sequence corresponding to codons 88 through 94 of the Fur sequence was used.

To produce ^{15}N- or ^{13}C-labeled protein, cells were grown in M9 medium supplemented with 40ug/mL kanamycin, 2uM ferric citrate, and 0.5g/L $^{15}NH_4Cl$ or 1g/L

Table 3.1 Amino acid sequence of the Fur protein

1	ala	asp	asn	asn	thr	ala	leu	lys	lys	ala
11	gly	leu	lys	val	thr	leu	pro	arg	leu	lys
21	ile	leu	glu	val	leu	gln	glu	pro	asp	asn
31	his	his	val	ser	ala	glu	asp	leu	tyr	lys
41	arg	leu	ile	asp	met	gly	glu	glu	ile	gly
51	leu	ala	thr	val	tyr	arg	val	leu	asn	gln
61	phe	asp	asp	ala	gly	ile	val	thr	arg	his
71	asn	phe	glu	gly	gly	lys	***ser***[a]	val	phe	***glu***[b]
81	leu	thr	gln	gln	his	his	his	***asp***[c]	his	leu
91	ile	cys	leu	asp	cys	gly	lys	val	ile	glu
101	phe	ser	asp	asp	ser	ile	glu	ala	arg	gln
111	arg	glu	ile	ala	ala	lys	his	gly	ile	arg
121	leu	thr	asn	his	ser	leu	tyr	leu	tyr	gly
131	his	cys	ala	glu	gly	asp	cys	arg	glu	asp
141	glu	his	ala	his	glu	gly	lys			

[a-c]Sequences of various metal binding domain constructs are indicated with the first residue of each domain italicized. [a]The tryptic fragment FurT starts with residue 77 (S77) of the Fur protein immediately following the K76 recognized by the protease. [b]Recombinant expressed metal binding domain T80 starts with a methionine (cloning artifact), followed immediately by residue 80 (E80) of Fur. [c]Recombinant expressed metal binding domain T88 starts with a methionine (cloning artifact), followed immediately by residue 88 (D88) of Fur. All purified recombinant metal binding domain constructs include the N-terminal methionine and all subsequent resdues through K147 of the Fur protein.

[^{13}C]-glucose. Proteins selectively labeled with ^{15}N-leucine and ^{15}N-histidine were grown similarly, but with the addition of the other 19 naturally occurring amino acids at a concentration of 100 mg/L, and the isotopically enriched amino acid at 150 mg/L. Cells were grown to an A_{600} of 0.38 and protein expression was induced by addition of 1mM IPTG, followed by growth for seven to eight hours. Protein was purified as described previously (Wee et al., 1988), except that a 40-80% ammonium sulfate saturation fraction was used as starting material for chromatography on zinc-iminodiacetate agarose. The identity of the FurT88 was confirmed by electrospray ionization mass spectroscopy.

Preparation of NMR Samples: Purified protein was prepared either by lyophilizing to dryness or by concentrating in a Centriprep-10 (Amicon) to the desired concentration. NMR samples ranged in protein concentration from 0.5 to 2 mM. Concentration was determined by measurement of absorption at 276 nm, using extinction coefficients calculated by the method of von Hippel (Gill & von Hippel, 1989). Samples were prepared in either 99.96% D_2O, or 10% D_2O/90% H_2O. Samples for structural NMR studies were buffered at pH 7.0 or 7.5 with either 20 mM phosphate buffer or 20 mM perdeuterated (D_{11}) Tris buffer in solutions containing from 0.5-2.0 mM dithiothreitol (DTT) to maintain the cysteine thiols in the reduced form. Samples for metal binding NMR studies were buffered in 20 mM D_{11}-Tris, pH 7.5 with 200 mM NaCl and 0.5 mM DTT. The non-denaturing detergent CHAPS (10-20 mM) was included in samples of FurT and T88 to obtain narrower resonances for multidimensional NMR experiments.

NMR Experiments: NMR experiments were performed at 500 MHz on a GE Ω500 NMR spectrometer or at 600 MHz on a Bruker AMX600. Experiments were collected on intact Fur and FurT at 35° C, while T88 and T80 were studied at 25° C. Chemical shift values were referenced externally to to TSP (^{1}H) and liquid ammonia (^{15}N) (Live et al., 1984). Non-acquisition dimensions of all multidimensional experiments utilized either time-proportional phase incrementation (TPPI) (Drobny et al., 1979) or the States-TPPI method for quadrature detection (Marion et al., 1989a). All data were processed with FELIX

version 2.30β (Biosym), including linear prediction calculations. Shifted skewed sine-bell and squared sine-bell functions were used for apodization of the free induction decays.

1D ^1H NMR spectra in D_2O were collected with a 1-pulse sequence with presaturation solvent suppression, and H_2O spectra used a 1-1 'jump-return' selective excitation pulse to suppress the intense solvent resonance. Most spectra used spectral widths of 6000-10 000 Hz, while metal binding experiments with Fur and T88 used wide spectral widths of 50 000 Hz to allow observation of hyperfine-shifted protein resonances.

2D NOESY (Anil-Kumar et al., 1980) and TOCSY (Bax & Davis, 1985) experiments were collected in both D_2O and H_2O solutions. The spectral width in all H_2O experiments was 10 000 Hz, and 7352 Hz in D_2O experiments. A total of 2048 data points were collected in t2 and 400-600 experiments were collected in t1, with 32 to 64 scans for each t1 increment. Solvent suppression in D_2O NOESY and all TOCSY experiments was achieved by low power presaturation of the water resonance for 1-2 s before each scan. NOESY experiments in H_2O used either presaturation or a 1-1 selective-excitation read pulse for solvent suppression.

^{15}N-^1H 2D 1-1 Echo HMQC (Sklenar & Bax, 1987) experiments were collected using the pulse sequence of Figure 1.6A. Similar spectral parameters to the 2D ^1H experiments were used in the ^1H dimension, and 128 or 256 complex points were acquired in the ^{15}N dimension with a spectral width of 3650 or 1861 Hz. The ^1H carrier was placed on the H_2O resonance at 4.80 ppm, and the ^{15}N carrier was set to 119 ppm. The ^{15}N-^1H HMQC experiment was used for purposes of assignment and comparison of the NH correlation footprint of the intact Fur protein and the metal binding domain constructs, and it also provided amide exchange information from a T88 sample dissolved in D_2O immediately prior to acquisition of a series of 2D experiments.

^{15}N-edited 3D NOESY-HMQC (Kay et al., 1989; Marion et al., 1989b), experiments were collected for FurT and T88 samples with spectral widths of 7352 Hz for the ^1H dimensions and 1861 Hz for the ^{15}N dimension. The 3D NOESY-HMQC

experiment collected on the FurT88 sample used the pulse sequence shown in Figure 1.8A, but a slightly different pulse sequence was used for the experiment collected on FurT. This version incorporated the 1-1 echo HMQC pulse sequence of Figure 1.6A to provide solvent suppresion without the use of presaturation. The NOESY mixing time was 160 ms for FurT and 80 ms for T88. An ^{15}N-edited 3D TOCSY-HSQC experiment was collected on FurT with a DIPSI spin-lock period for isotropic mixing of 60 ms. This experiment differs from the pulse program described in Figure 1.8B in that the separation by ^{15}N chemical shift is performed by an HSQC instead of HMQC to allow water suppression by homospoil and spinlock pulses instead of presaturation. A total of 128 x 24 x 1024 complex points were collected in the t1, t2, and t3 dimensions, respectively. Data were apodized in each dimension with a skewed shifted sinebell function. A shift of 75° was used in each dimension, with a skew of 1.0, 0.8, and 0.5 in the t1, t2, and t3 dimensions, respectively. Data were zero-filled to yield a 512 x 64 x 512 real matrix upon Fourier transformation. The 3D ^{15}N-edited 3D NOESY-HMQC and TOCSY-HMQC pulse sequences shown in Figure 1.8 were also used in 2D mode, without incrementation of the t2 period, to acquire ^{15}N half-filtered NOESY and TOCSY spectra of the selectively labeled FurT samples. Pulse programs for all 2D and 3D experiments used can be found in Appendix A.

NMR studies of intact Fur protein

The optimistic goal of NMR studies on the Fur protein was two-fold. First, using standard techniques for determination of the solution structure of proteins, elucidate the three-dimensional fold of the Fur protein. Second, characterize the location and mode of metal-binding in the context of this structure, by direct observation of metal-protein interactions, if possible. It became apparent that this was a much more formidable task than initially perceived, due primarily to the fact that the Fur protein is a dimer of 35 kDa in solution, rather than a monomer of 17 kDa. Gel filtration studies at reasonably high

concentration (but still lower than typical NMR concentrations) indicated the presence of tetrameric species also in equilibrium with the dimer.

Initial approaches to obtaining NMR spectra of the Fur protein were thwarted by the difficulties inherent in obtaining high-quality spectra of the exchangeable NH protons of such a large protein, and the additional complication of rapid exchange with solvent at neutral pH. Because the Fur protein precipitated at pH below 7.0 at NMR concentrations, all NMR experiments were collected at pH 7 or higher. Spectra acquired with presaturation of the H_2O resonance resulted in significant reduction in the intensity of many resonances, especially in the region from 7-10 ppm. This effect is due to a combination of two factors. First, exchange of NH protons with the solvent during the preirradiation period results in the partially or completely saturated magnetization of the solvent resonance being substituted for the normally unaffected resonance of the protein. Additionally, any resonance of the protein may be affected by spin-diffusion from nearby spins which are saturated by the presaturation RF field. If a given αH resonates at or near the chemical shift of the water, it may be saturated by the presaturation pulse and, via cross-relaxation, can reduce the intensity of any other spins within the normal range of NOE interactions (i.e., < 5Å).

The problems associated with presaturation methods can be alleviated in some cases with the use of selective excitation, shaped pulses, or spin-lock pulses for solvent suppression, but none of these is generally applicable to all types of multidimensional experiments. The advent of pulsed field gradient (PFG) technology represents the best solution to these problems, and others as well.

Another significant problem was the difficulty in obtaining 2D NMR data from the basic NOESY and TOCSY experiments described in Chapter 1 and applied in Chapter 2 to a small peptide. The overriding factor in this instance is the short ^1H relaxation times, T_2, of such a large protein. As the size of a molecule and, consequently, its rate of reorientation in solution increase, the average T_2 relaxation time of protons in the molecule

decreases. This is manifested in the broadening of resonances in the NMR spectrum, and reduces the number of magnetization transfer steps which may be utilized in a multidimensional NMR experiment.

The ^{15}N-^{1}H HMQC spectrum of Fur is shown in Figure 3.1. While a spectrum of this sort may be collected with little difficulty with the pulse sequence of Figure 1.6A, a spectrum collected using the CBCA(CO)NH pulse sequence of Figure 1.10B would likely contain little or no signal. While the total duration of the pulse sequence of Figure 1.6A in this case is ~9 ms plus the value of t1, the duration of the CBCA(CO)NH pulse program would be over 60 ms, which may be several times the value of T_2, resulting in effectively no magnetization remaining to be detected at the completion of the pulse sequence. Even the basic 2D TOCSY experiment of Figure 1.3B will suffer the same limitation from rapid relaxation during the 40-80 ms isotropic mixing period.

Because of the clear limits to collecting multidimensional NMR data on Fur, even with the incorporation of heteronuclei like ^{15}N, an alternative to studying the entire Fur protein was desirable. Recent studies of the Fur protein using limited proteolysis to probe the existence and location of subdomain structure (Coy & Neilands, 1991) provided this opportunity.

NMR studies of FurT and T88

Resonance assignments of FurT: Because the intact Fur protein had unfavorably large linewidths and poor relaxation characteristics due to its large size and multimeric state (35 kDa dimer or 70 kDa tetramer at NMR concentrations), an attempt to study a subunit of the Fur protein known to contain the essential metal-binding elements was begun. Initial comparison of the 1D spectra of Fur and FurT reveals general similarity in the amount of chemical shift dispersion, but no definition of the true structural similarity can be made from such a comparison. A more informative analysis can be made from the 2D ^{15}N-^{1}H

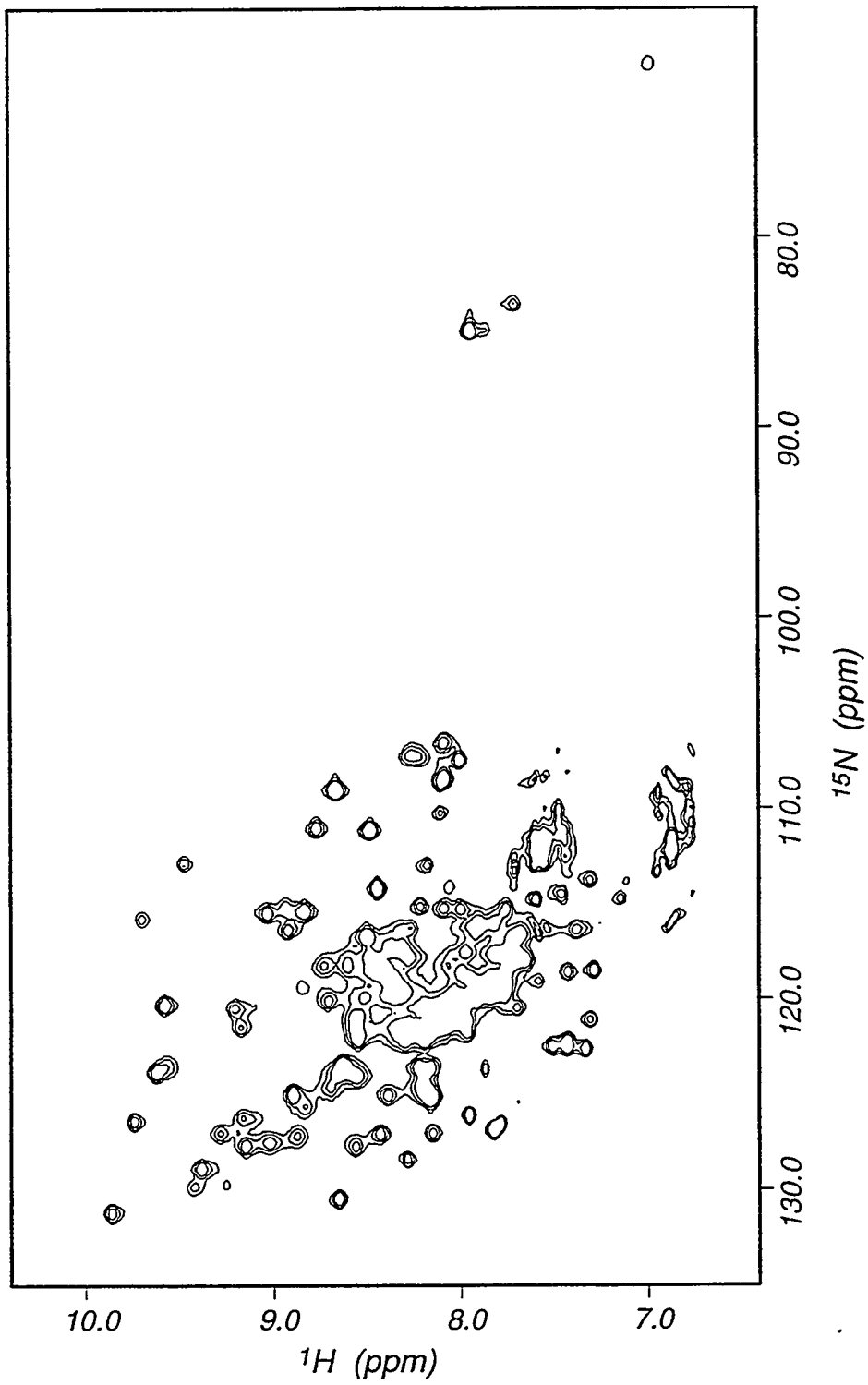


Figure 3.1: ^{15}N-^{1}H HMQC spectrum of the full-length Fur protein. The 1-1 echo HMQC experiment described in Chapter 1 (Figure 1.6A) was used to collect a total of 512 t_1 points with an ^{15}N spectral width of 7300 Hz. Some arginine and lysine sidechain NH's can be seen at upfield ^{15}N chemical shift values.

HMQC spectra of the two proteins, as shown in Figure 3.2. Two aspects of the pair of spectra are immediately obvious. One is the strong similarity in the location of many of the crosspeaks in both spectra, indicating that truncation of Fur does not destroy the folded structure of the C-terminal domain. Also apparent is difference in linewidths of the two proteins in both the ^{15}N and ^{1}H dimensions, consistent with the nearly twofold difference in their molecular weights. This narrower linewidth was the primary goal of studying a smaller portion of the Fur protein.

2D NMR of FurT The advantage of this reduced linewidth is the ability to successfully apply some of the 2D NMR experiments to FurT which were not feasible with the intact Fur protein. A region of the 2D NOESY spectrum of FurT is shown in Figure 3.3A. Although the improvement in nuclear relaxation times was dramatic with FurT compared to Fur, the non-denaturing detergent CHAPS was added to samples to produce even narrower resonances. The use of CHAPS as a way to enhance solubility or reduce non-specific association of proteins for improved NMR spectral properties had been described in the literature (Anglister et al., 1993), and gave reproducible improvements in the dispersion and linewidth characteristics of the FurT samples used.

The primary drawback to using CHAPS was the presence of very intense lines in the proton spectrum from the detergent molecules at 10-20 mM. The difficulty is eliminated by using heteronuclear labeling and eliminating signals from protons which are not correlated to a heteronucleus. This is illustrated in Figure 3.3 by comparison of the NOESY spectrum of unlabeled protein in panel A with the ^{15}N-edited NOESY spectra of panels B and C. The spectra of Figure 3.3B and C were collected on FurT samples which had been selectively labeled with ^{15}N-His and ^{15}N-Leu, respectively, so that these experiments contain crosspeaks due only to NOE interactions between the backbone amide protons of these residues (F2 axis) and the other protons of the molecule which are within 5Å. These two spectra are aqcuired with a pulse program which is essentially identical to the 3D ^{15}N-separated NOESY-HMQC of Figure 1.8A, except that the t2/2 evolution

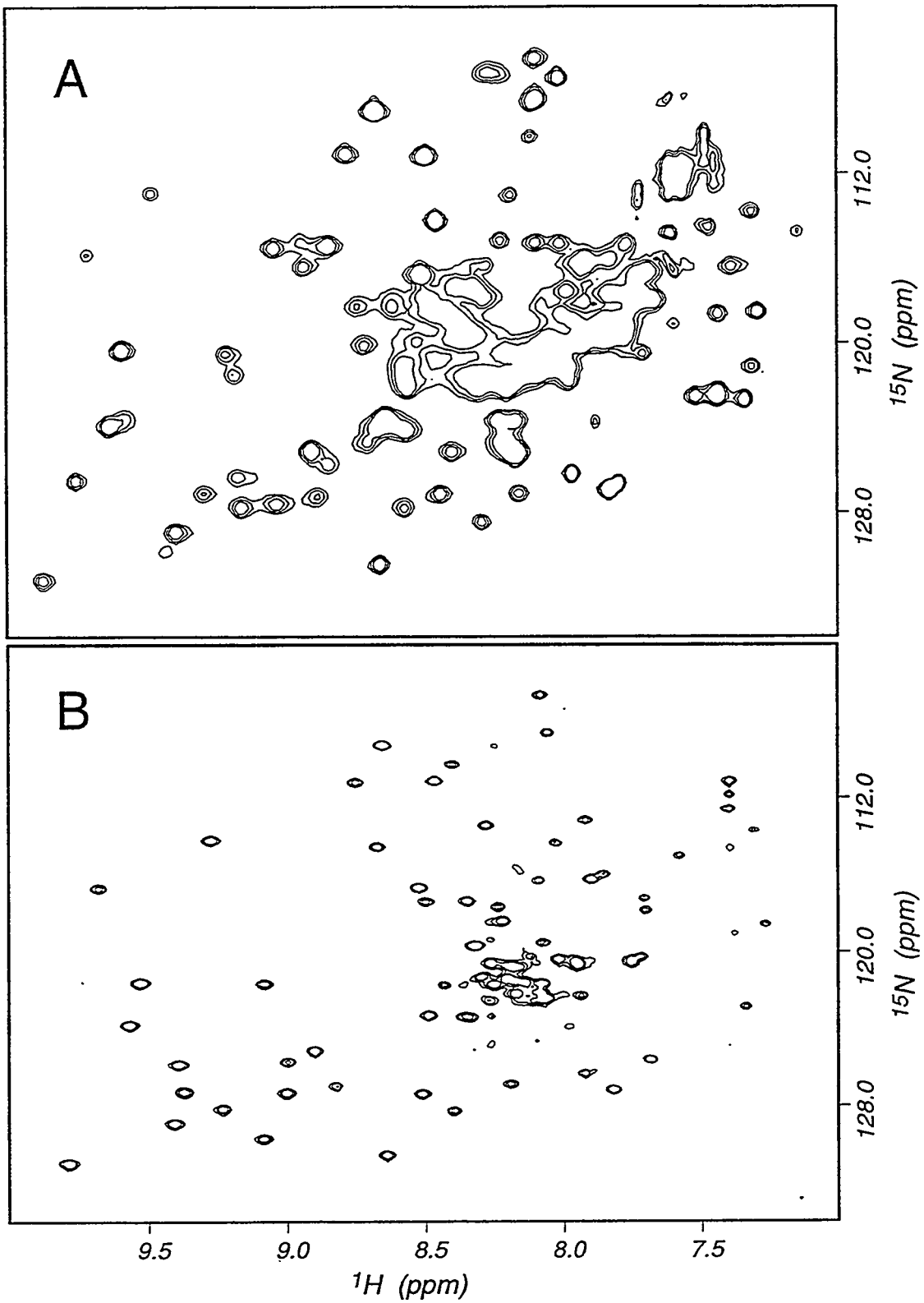


Figure 3.2: ^{15}N-^{1}H HMQC spectra of the full-length Fur protein (A) and the trypic fragment of Fur, FurT (B). The spectrum of panel A was collected with a larger ^{15}N spectral width than that of panel B, resulting in slightly lower digital resolution in that dimension of the Fur spectrum.

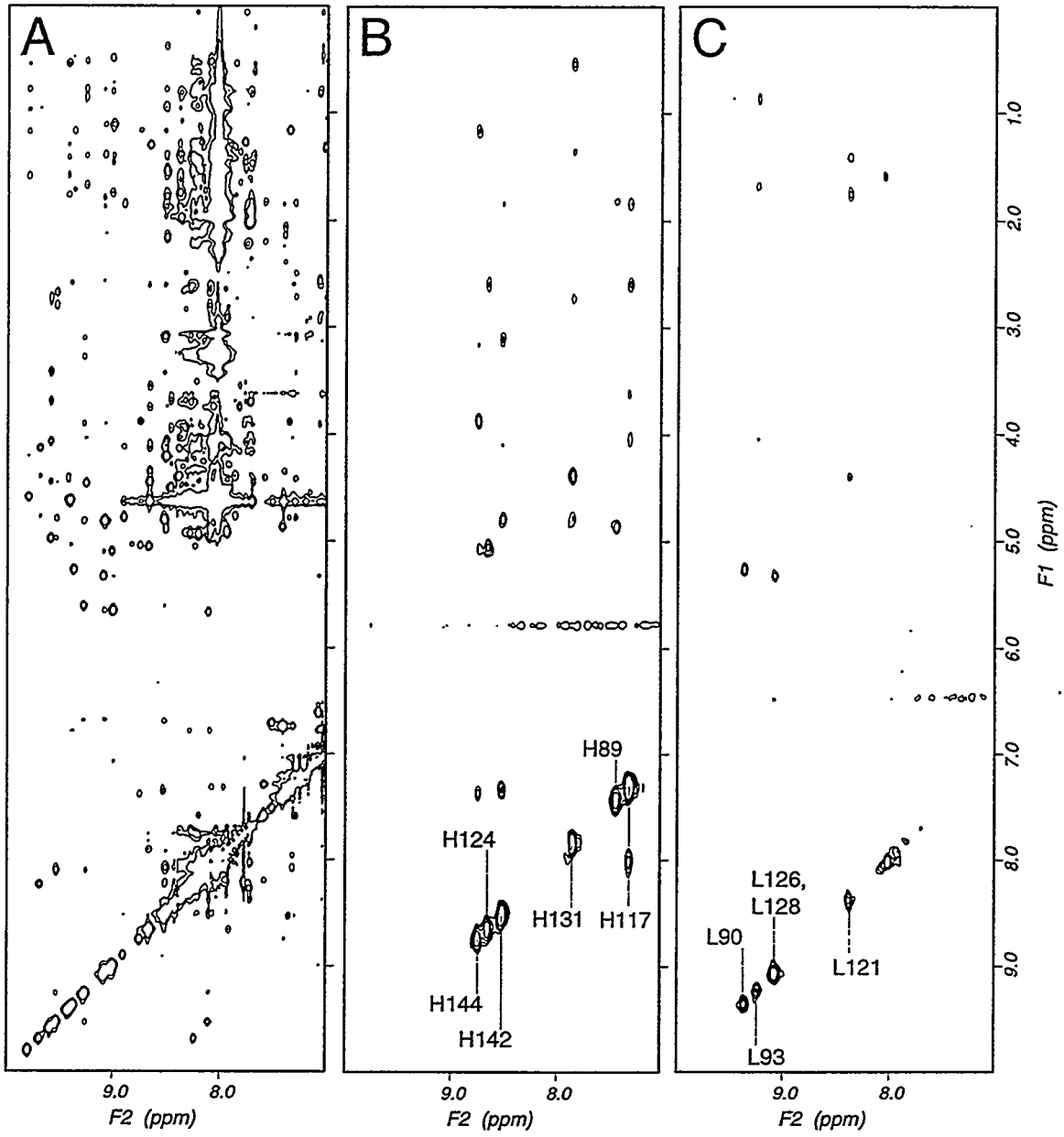


Figure 3.3: NOESY spectra of FurT. A. 2D 125 ms NOESY of unlabeled FurT. B. 2D ^{15}N-edited 125 ms NOESY of ^{15}N-His-labeled FurT. C. 2D ^{15}N-edited 75 ms NOESY of ^{15}N-Leu-labeled FurT. All spectra were collected at 35° C, with 20 mM CHAPS non-denaturing detergent. ^{15}N-edited NOESY spectra were acquired using a modified version of the 3D ^{15}N-separated NOESY-HMQC pulse sequence shown in Figure 1.8A, which incorporated the 1-1 echo HMQC sequence for solvent suppression. The 2D NOESY of panel A also used a 1-1 read pulse for solvent suppression. All signals in the F2 dimension which arise from non-^{15}N-bound protons are suppressed in the ^{15}N-edited NOESY spectra of panels B and C. This includes the intense amide resonance of the 20 mM CHAPS detergent, clearly visible in panel A as large contours at F2 = 8 ppm.

periods are set to zero and not incremented. The resulting dataset has two ^{1}H dimensions, F1, which is not ^{15}N-edited and can contain any ^{1}H resonances of the normal spectrum, and F2, which is ^{15}N-edited, filtering out any signals from protons which are not directly bound to an ^{15}N atom.

In a similar fashion, ^{15}N-edited TOCSY spectra of these selectively labeled samples can be collected. Figure 3.4 shows the ^{15}N-edited NOESY and TOCSY spectra of the ^{15}N-His FurT sample with a number of the sequential NOE crosspeaks labeled with the connecting residue. With a large variety of selectively labeled NMR samples, significant sequential assignment information can be derived with these ^{15}N-edited 2D methods, but in some cases it becomes necessary to use the ^{15}N chemical shift to provide additional dispersion of spin systems.

While the edited 2D NOESY spectra of Figure 3.3(B, C) are sparser than the regular NOESY (Figure 3.3A), they still rely on the dispersion of only the ^{1}H chemical shifts. Figures 3.5 and 3.6 show the HMQC spectra of the ^{15}N-His-labeled and ^{15}N-Leu-labeled FurT proteins, in which the NH resonances of the histidines and leucines are spread in two dimensions by both the ^{1}H and ^{15}N chemical shifts. While the histidine amide resonances are completely resolved in the ^{1}H dimension, the amide proton chemical shifts of two of the leucine residues are identical, resulting in the degeneracy of crosspeaks to the NH of L126 and L128 in both the normal 2D NOESY and the ^{15}N-edited NOESY. This ambiguity is easily removed by the addition of an ^{15}N frequency dimension, as shown in the HMQC of ^{15}N-Leu FurT (Figure 3.6) in which the L126 and L128 resonances are widely separated by their respective ^{15}N chemical shifts. This example serves to illustrate the need to use heteronuclear labels not only for editing of 2D spectra but also for separation of resonances into a third dimension.

3D NMR of FurT: To take full advantage of the improved dispersion of 3D ^{15}N-separated experiments, a uniformly ^{15}N-labeled FurT sample was prepared and a 3D NOESY-HMQC spectrum was collected. This spectrum, in combination with the

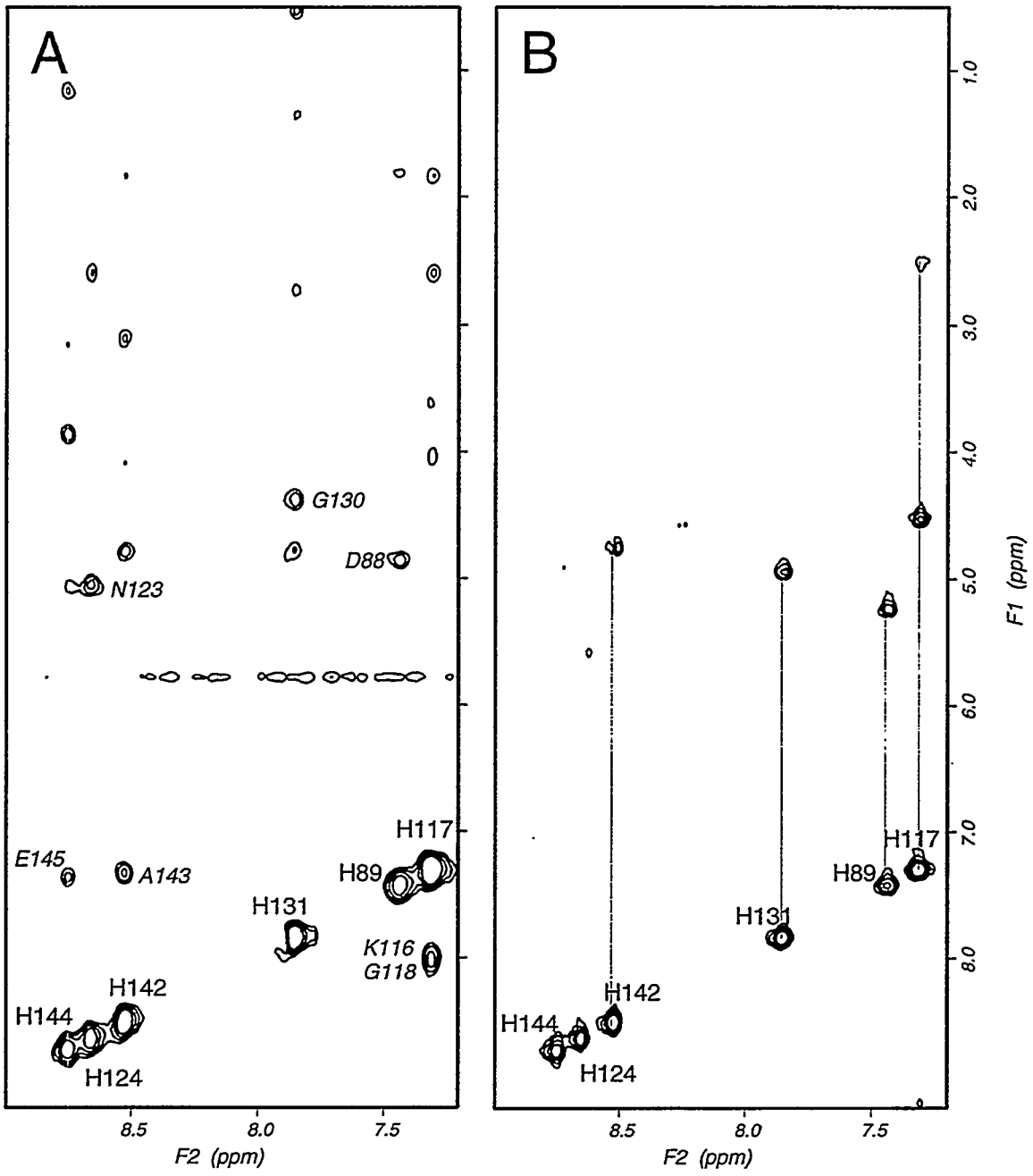


Figure 3.4: ^{15}N-edited NOESY (A) and TOCSY (B) spectra of ^{15}N-His-labeled FurT collected at 35° C with 20 mM CHAPS. Assignments are indicated with one-letter amino acid abbreviation and residue number. Crosspeaks which appear in the NOESY but not the TOCSY may be interresidual correlations which can be used for sequential assignment purposes. Some sequential NOEs are indicated with residue number and type. Because of large proton linewidths, a short isotropic mixing period was required in the ^{15}N-edited TOCSY, resulting in few observable correlations between backbone amide and sidechain protons. Vertical lines connect diagonal peaks with intraresidue αH and sidechain crosspeaks in the TOCSY spectrum.

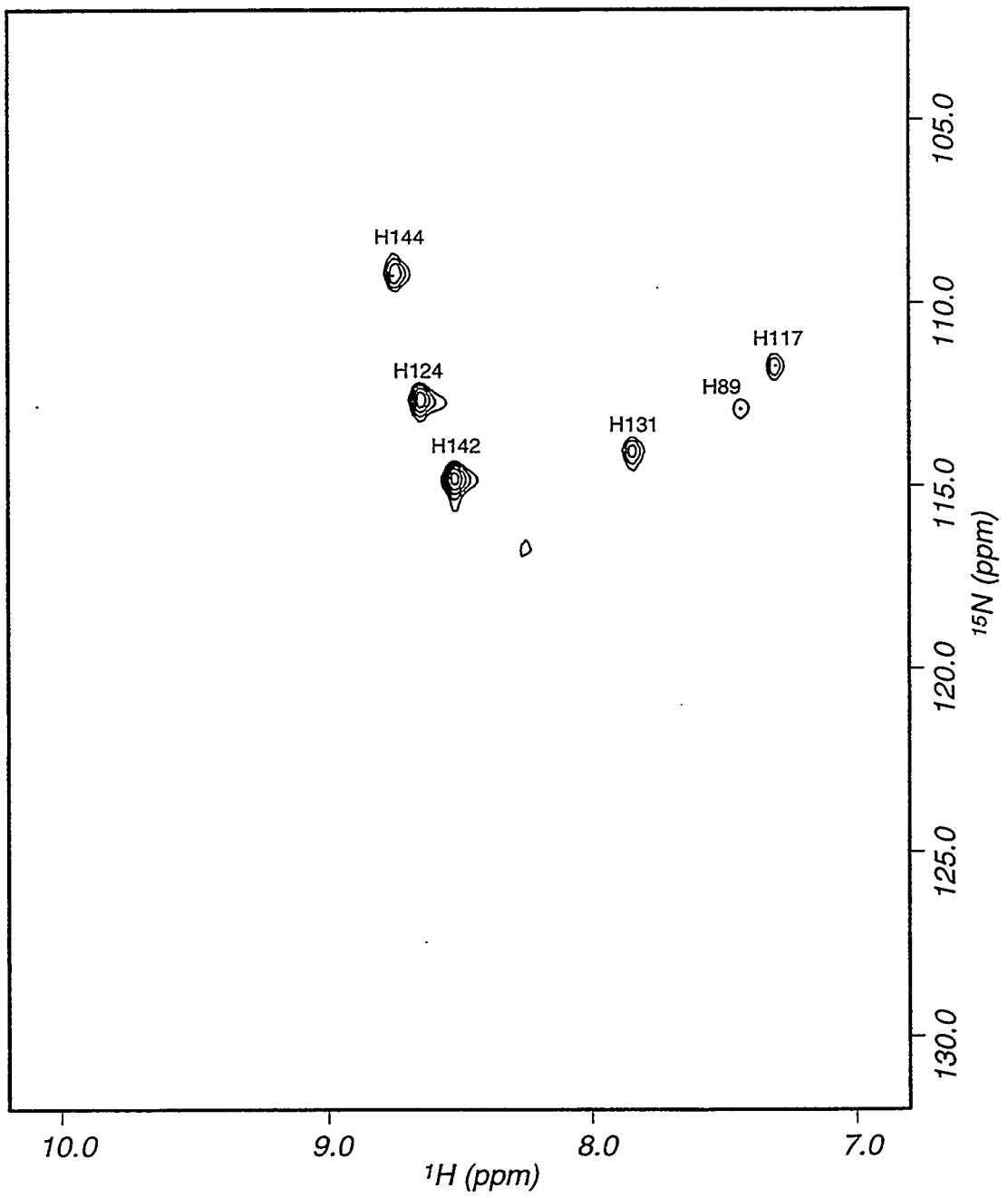


Figure 3.5: ^{15}N-^{1}H HMQC spectrum of ^{15}N-His-labeled FurT collected at 35° C, with 20 mM CHAPS. Assignments are indicated with one-letter amino acid abbreviation and residue number. Resonances observed in this spectrum are a subset of the spectrum of uniformly ^{15}N-labeled FurT. N-terminal residues of FurT appear to be dynamically disordered, and H85, H86 and H87 do not give rise to strong signals in this spectrum, despite being ^{15}N-labeled.

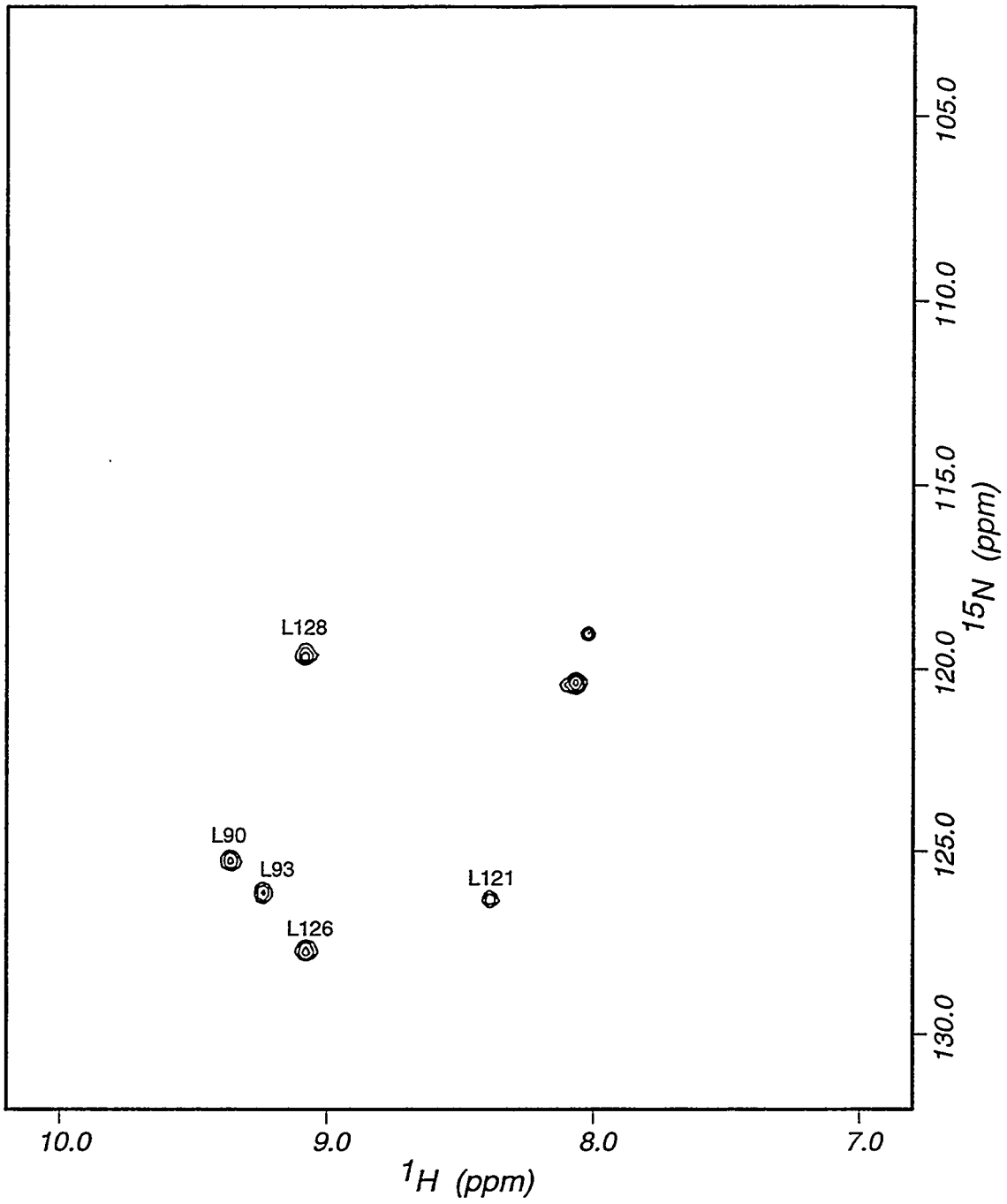


Figure 3.6: ^{15}N-^{1}H HMQC spectrum of ^{15}N-Leu-labeled FurT collected at 35° C, with 20 mM CHAPS. Assignments are indicated with one-letter amino acid abbreviation and residue number. Resonances observed in this spectrum are a subset of the spectrum of uniformly ^{15}N-labeled FurT. The HMQC spectrum clearly resolves the NH resonances of L126 and L128 which are degenerate in ^{1}H chemical shift but distinct in ^{15}N chemical shift. Unlabeled peaks could not be unambiguously assigned to specific leucine residues.

information derived from the selectively labeled samples, allowed the sequential assignment of over half of the residues of FurT. The manner in which these assignments were obtained will be described briefly.

The HMQC (Figure 3.2B) and 2D NOESY spectra (Figure 3.3) of FurT can be used to describe the approach used to analyze the 3D NOESY-HMQC spectrum. This 3D spectrum contains essentially all of the information in the region of the 2D NOESY shown in Figure 3.3A, but in the 3D experiment it is distributed orthogonally from that ^{1}H-^{1}H plane according to the pattern of ^{15}N chemical shifts shown in the HMQC of Figure 3.2B. In other words, if the sum of all signals in the 3D spectrum were taken along the ^{15}N dimension, the resulting projection would be identical to the 2D NOESY, while the analagous projection taken along the second ^{1}H dimension would produce the HMQC spectrum.

Because the HMQC spectrum of Figure 3.2B specifies the location of the set of NOESY crosspeaks to a given amide proton, it can be helpful in extracting all such sets of crosspeaks from the 3D spectrum by pulling narrow strips at the appropriate ^{1}H (F3) and ^{15}N (F1) chemical shifts and placing them into a single two-dimensional array. This method of using 'strip plots' for analysis of 3D data is illustrated in Figure 3.7. A set of strips from the 3D NOESY-HMQC spectrum of FurT88 are placed in order of the sequence from residues E139-G146. This strip plot shows the ability of the ^{15}N-dimension of the 3D experiment to separate degenerate amide proton resonances. The amide protons of D140, A143, and E145 all have very similar chemical shifts, and the three strips shown in Figure 3.7 for these residues would normally be superimposed on each other in a 2D NOESY spectrum, resulting in ambiguity in analyzing which of the crosspeaks at that amide proton chemical shift are correlated. Separation of amide ^{1}H resonances by the amide ^{15}N chemical shift in the 3D NOESY-HMQC has clearly eliminated that ambiguity.

The assignment of the residues shown in Figure 3.7 was obviously facilitated by the knowledge that two of these strips are identified as histidines by comparison of the

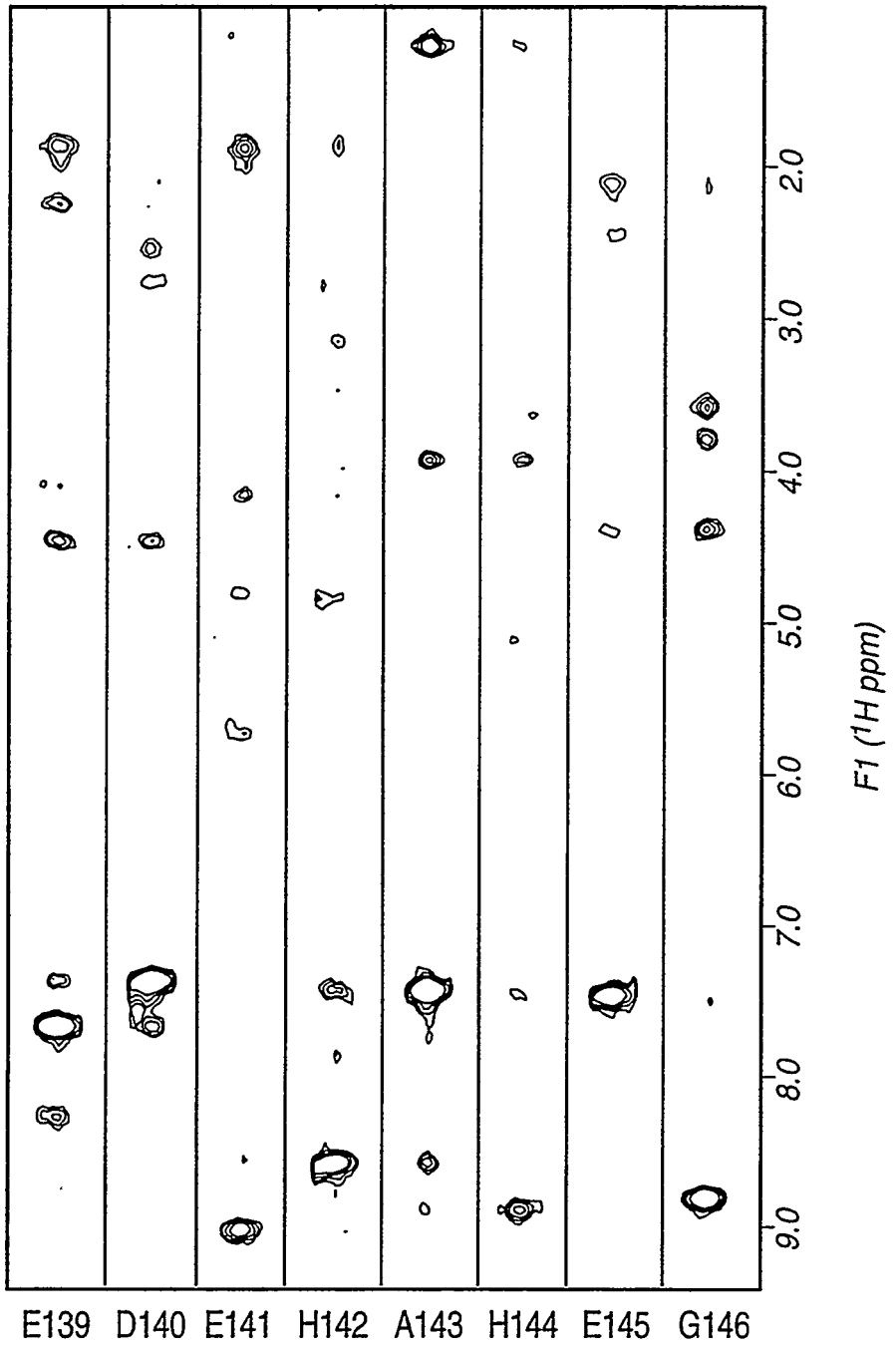


Figure 3.7: Strip plot from the 3D ^{15}N-separated NOESY-HMQC spectrum of FurT88. Each strip is centered along a vector from the 3D matrix which is determined by the amide ^{15}N and ^{1}H chemical shifts of the residue indicated on the horizontal axis. Comparison of this plot with the NOESY spectrum of Figure 3.3A illustrates the high degree of separation of NOESY crosspeaks obtained with this 3D NMR experiment. Some of the crosspeaks which connect this sequence of strips can be seen, such as the NH-NH crosspeaks from E139-D140 and H142-A143.

HMQC spectra of ^{15}N-His FurT and uniformly ^{15}N-labeled FurT. Complete analysis of the 3D NOESY-HMQC spectrum of FurT resulted in the assignment of a majority of amide ^{15}N and ^{1}H resonances as shown in the HMQC spectrum of Figure 3.8. Sequential assignments obtained for FurT are also listed in Table 3.2. Assignments for approximately 10 residues at the N-terminus of the FurT domain were not obtained. Dynamic mobility of these residues reduces the intensity of resonances as well as the rate of NOE buildup. Other residues could not be unambiguously assigned due to overlap in the 3D spectra and a lack of intraresidue correlations providing identification of amino acid types.

The primary barrier to completing the assignments of FurT arises from the inability to obtain high-quality 3D TOCSY-HMQC spectra of the FurT protein. Despite the improvement of FurT relative to intact Fur, the ^{1}H linewidths of FurT still restrict the amount of TOCSY-type information which can be generated. A number of NH-αH correlations were obtained from the 2D TOCSY spectrum of FurT in the presence of CHAPS, and these were used to assist the correlation of strips of the 3D NOESY-HMQC spectrum. Obtaining complete intraresidue correlation information continues to be the most difficult aspect of the ongoing NMR studies of the metal-binding domain of Fur.

Resonance assignments and secondary structure of FurT88: When the metal binding domain of Fur was subcloned and expressed as a domain independent of Fur, its NMR spectrum was initially examined to confirm its similarity to that of FurT. Figure 3.9 shows the HMQC spectra of FurT88 and FurT, and close inspection reveals them to be nearly identical. Many of the assignments made with FurT could be transferred to FurT88 based just on the similarity of the HMQC spectra, but additional NMR experiments were performed on uniformly ^{15}N-labeled FurT88 protein to confirm and augment the FurT assignments.

3D ^{15}N-separated NOESY-HMQC and TOCSY-HSQC spectra were collected on FurT88 at 25° C, rather than at 35° C as was used for the FurT studies. FurT88 samples contained 20 mM CHAPS as was used in experiments with FurT. While the 3D NOESY-

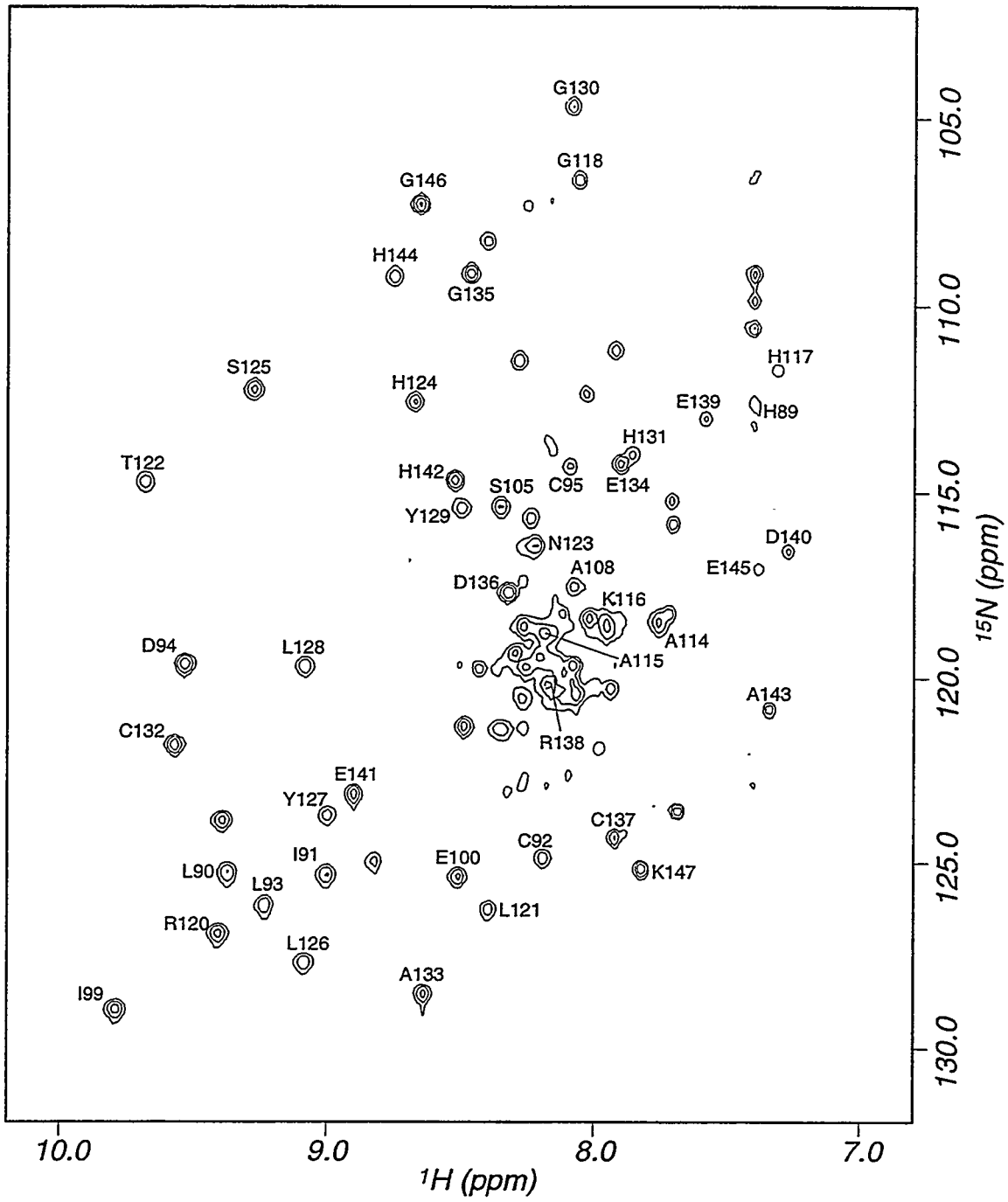


Figure 3.8: ^{15}N-^{1}H HMQC (1-1 echo HMQC) spectrum of Fur(77-147) (FurT). Sequence-specific assignments are indicated with residue number and one-letter amino acid code. Numbering is that of the sequence of the full-length Fur protein. Spectrum was collected at 600 MHz at 35° C. Solution contained 20 mM phosphate buffer (pH 7.0) and 20 mM CHAPS.

Table 3.2 ^{1}H and ^{15}N resonance assignments of FurT, the metal-binding domain of the Fur protein

Residue	^{15}N	NH	αH	βH	other
Asp 88			4.85		
His 89	114.8	7.44	5.25	2.66, 2.53	
Leu 90	125.28	9.38	5.10	1.75, 1.6	1.2
Ile 91	125.37	9.00	4.90	1.65, 1.69	
Cys 92	124.87	8.20	4.45	3.1	
Leu 93	126.14	9.23	4.00	1.65,1.35	
Asp 94	119.59	9.53	4.80	2.82	2.69
Cys 95	114.28	8.08	4.90	3.15	3.30
Val 98	115.35	8.35	4.55		
Ile 99	128.95	9.79	4.02	1.60	
Glu 100	125.40	8.51	4.85	1.75	
Ala 114	118.5	7.76	4.15		
Ala 115	118.75	8.19			
Lys 116	118.55	7.96	4.03		
His 117	111.70	7.31	4.55	2.56,2.50	
Gly 118	106.62	8.06	4.10		
Leu 121	126.25	8.40			
Thr 122	114.71	9.68	4.12	0.90?	
Asn 123	116.41	8.22	5.05	2.51,2.62	
His 124	112.57	8.68	5.60	3.10, 3.28	
Ser 125	112.23	9.28	4.80	3.90	
Leu 126	127.72	9.08	5.65	2.10, 1.87	1.37
Tyr 127	123.74	9.00	5.32	2.93, 2.53	
Leu 128	119.69	9.08	4.80		
Tyr 129	115.37	8.51	5.67	2.79, 2.59	
Gly 130	104.67	8.09			
His 131	113.97	7.86	4.95	2.70	
Cys 132	121.81	9.57	4.45		
Ala 133	128.56	8.64	4.14	1.31	
Glu 134	114.22	7.90	4.49	1.85	
Gly 135	109.14	8.47	3.90, 3.68		
Asp 136	117.69	8.33	4.90	2.80	
Cys 137	124.32	7.92	4.40	2.88	
Arg 138	120.17	8.17	4.15		
Glu 139	113.01	7.57	4.38	1.85	2.12, 2.22
Asp 140	116.54	7.30	4.78	2.52, 2.72	
Glu 141	123.13	8.89	4.12	1.85, 1.95	
His 142	114.66	8.52	4.76		
Ala 143	120.83	7.32	3.88	1.20	
His 144	109.21	8.75	5.04	3.55, 3.18	
Glu 145	117.01	7.34	4.35	2.07, 2.13	2.40
Gly 146	107.27	8.65	3.76,3.56		
Lys 147	125.16	7.82	4.19	1.67,1.82	1.33

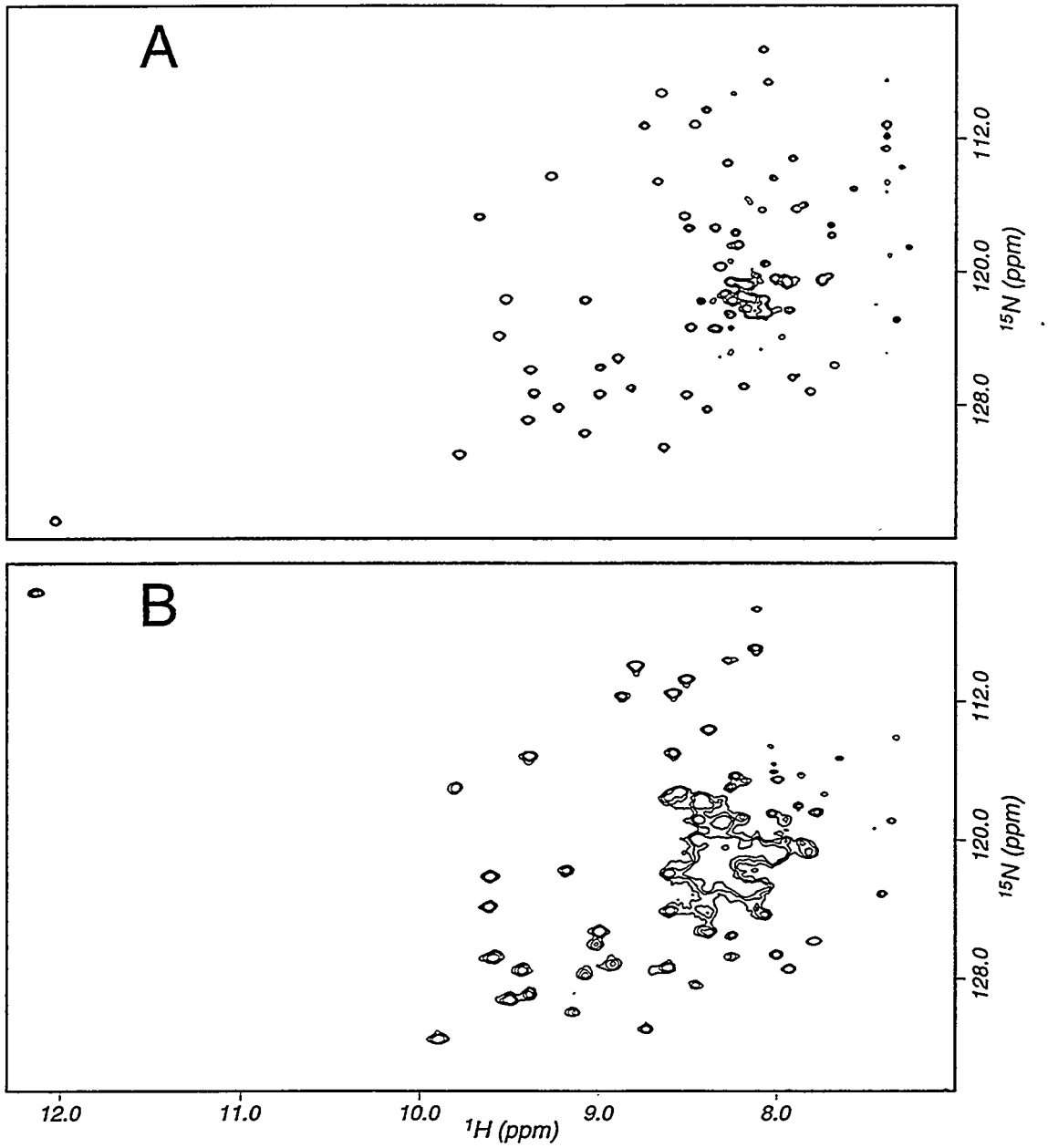


Figure 3.9: ^{15}N-^{1}H HMQC (1-1 echo HMQC) spectra of Fur(77-147) (FurT) (A) and Fur(88-147) (FurT88) (B). The FurT sample was produced by proteolysis of intact Fur protein with trypsin, while the FurT88 sample was expressed as a folded domain independent of the N-terminal half of Fur. The sequences of Fur and the C-terminal domain constructs are shown in Table 3.1. Inspection of the HMQC 'footprint' of the two proteins indicates that the tryptic fragment and the recombinant domain have essentially the same folded structure. This is also supported by the appearance of the histidine imidazole NH at 12 ppm in the spectra of both samples, which must be involved in hydrogen bonding and/or hydrophobic packing interactions to be observable. This resonance appears at different ^{15}N chemical shifts in the two spectra due to differences in spectral widths and the folding properties of different quadrature detection schemes.

HMQC spectrum of FurT88 (see Figure 3.7) provided useful information, the 3D TOCSY-HSQC experiment contained crosspeaks to very few of the backbone amides. The amide ^{15}N and ^{1}H assignments for FurT88 at 25° C and pH 7.5 are indicated in the HMQC spectrum shown in Figure 3.10. A table of assignments for FurT88 at these conditions is also included in Appendix C. HMQC experiments collected on FurT88 which is dissolved in D_2O were used to identify amides which exchange slowly with solvent, indicating the likely presence of secondary structure elements in those regions of the sequence.

Figure 3.11 shows a table of sequential NOEs (panel A) and a model of the β-sheet topology (panel B) for the metal binding domain of Fur, based on the 2D NOESY and 3D NOESY-HMQC spectra of FurT88. The indications of β-strand and helix shown in the NOE table are consistent with the patterns of sequential NOEs, but have not been confirmed with medium- or long-range NOEs. Also indicated in Figure 3.11A are the backbone amides which have reduced solvent exchange rates. The β-sheet diagram of Figure 3.11B is simply a model which is consistent with the sequential NOEs, a few long-range αH-αH crosspeaks and the amide exchange data. It indicates a possible scheme for the symmetric dimerization of Fur via an intermolecular β-sheet structure, represented with a dotted line at the interface of the monomer subunits. It is hoped that future studies will result in the refinement and expansion of this model. One aspect of the proposed secondary structure which is worth noting is the proximity of three of the four cysteine residues of the metal-binding domain. The presence of some type of turn in the backbone conformation of residues 93-95 is indicated by the sequential NH-NH crosspeaks, and would likely place the sidechains of C92 and C95 closer than if those residues were fully extended. The possiblity that these cysteine sidechains are part of the metal binding site will be explored in experiments planned for the near future.

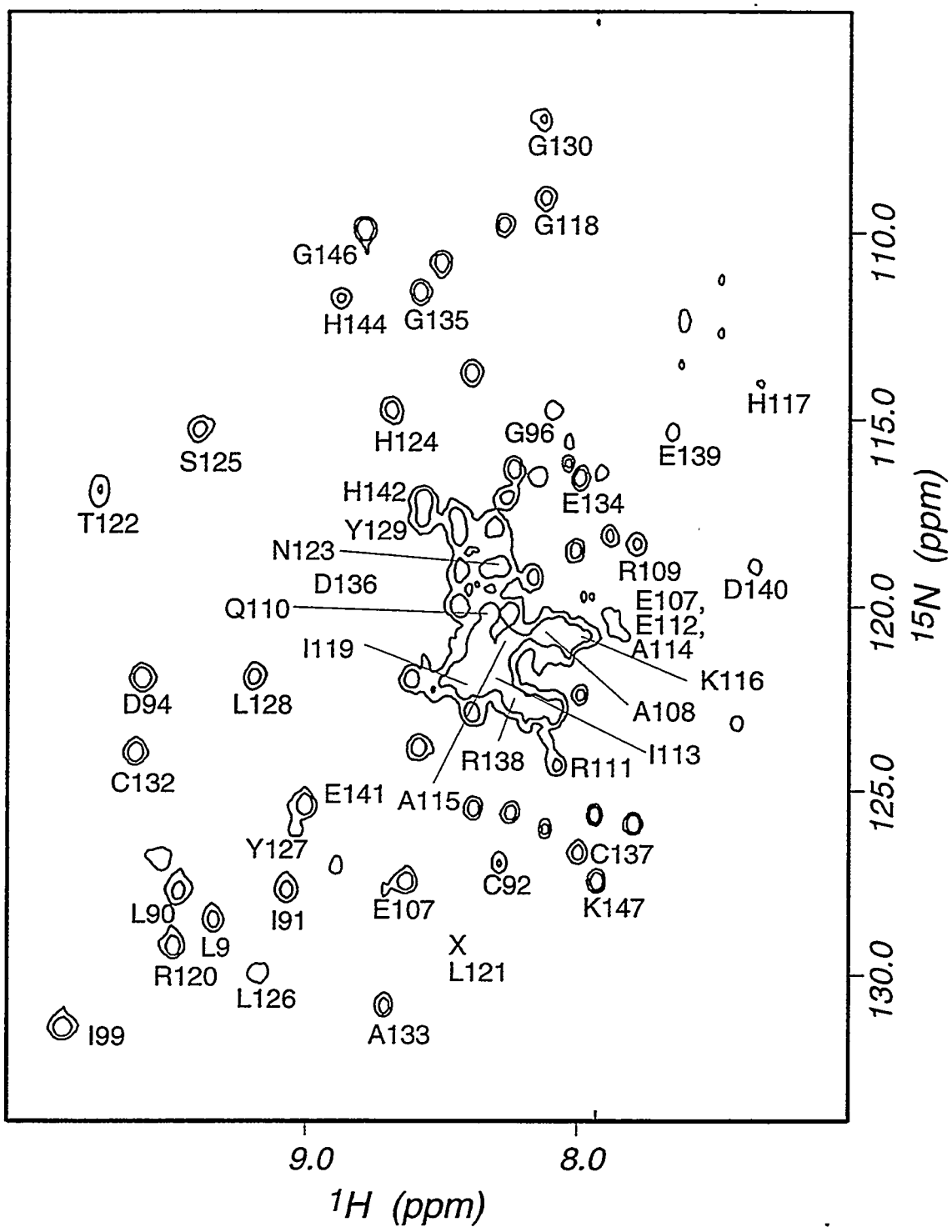


Figure 3.10: ^{15}N-^{1}H HMQC spectrum of FurT88, collected at 25° C, pH 7.5, with 20 mM CHAPS. Assignments are indicated with the one-letter amino acid code and residue number. No dramatic chemical shift differences between this spectrum and that of FurT are observed.

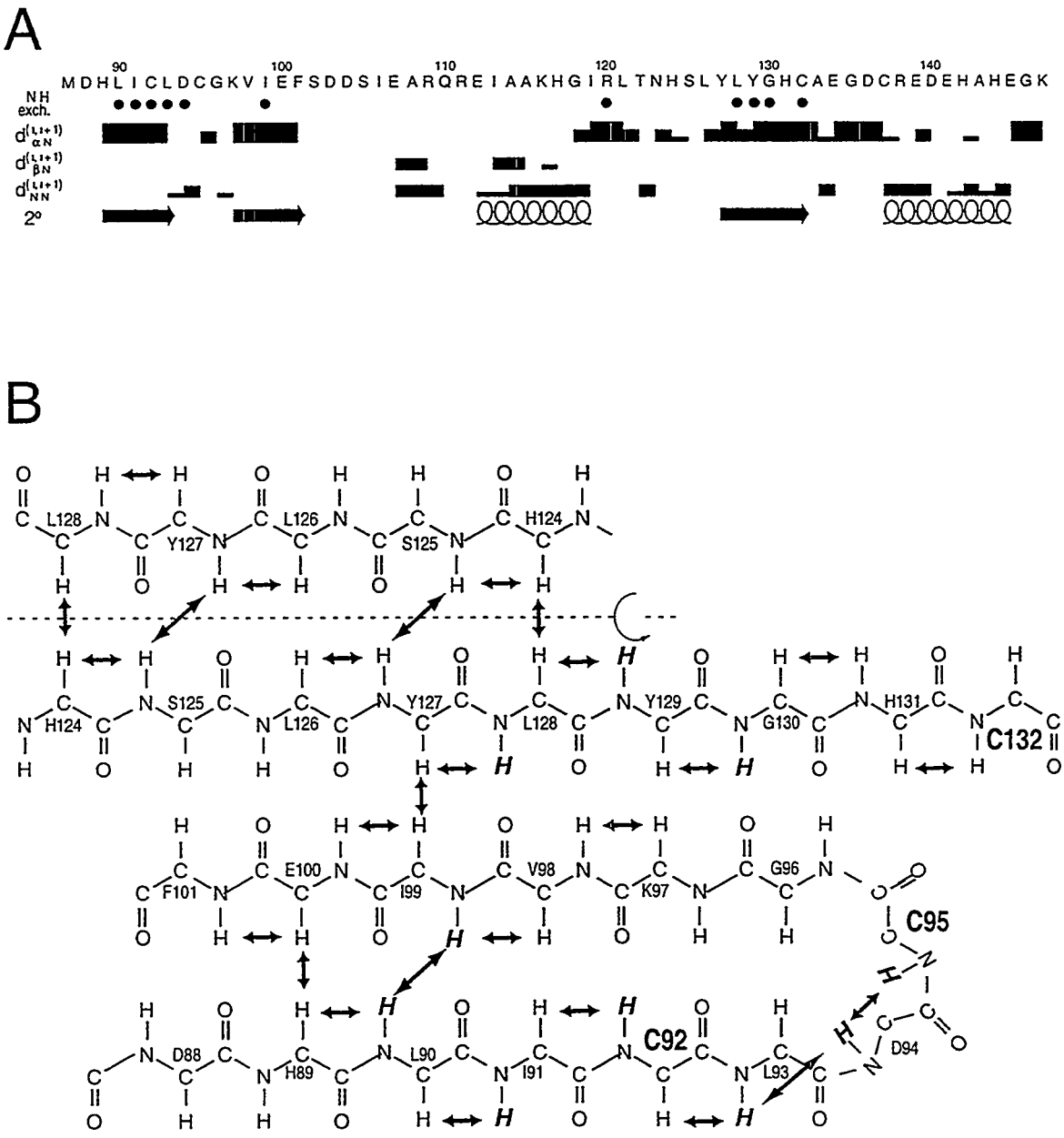


Figure 3.11: Secondary structure of the metal-binding domain of Fur. A. Sequential NOE diagram with regions of the sequence labeled as helix or sheet, based on the patterns of observed sequential $d_{\alpha N}$ or d_{NN} connectivities, respectively. Slowly exchanging amides are also indicated B. Model of β-sheet topology of metal-binding domain of the Fur protein. Regions of the sequence shown have reduced amide proton exchange rates (slowly exchanging NH's highlighted). Sequential and cross-strand NOEs are indicated with arrows. Three of the four cysteine residues (large boldface residue labels) are near each other in this model.

Metal binding studies of Fur and T88

To confirm that the Fur protein retains the ability to bind metal under the conditions used for structural NMR analysis, and to probe the types of interactions which might be involved in metal binding, a number of experiments with various divalent transition metal ions were attempted. Similarly, NMR spectra of the T88 metal binding domain construct were acquired in the presence and absence of metal to look for effects similar to those observed with intact Fur, to show that the metal binding sites were unchanged in the smaller protein.

A set of experiments to look for interactions of the histidine residues of FurT were performed using varying levels of Mn^{2+}, which is an activating metal of Fur, i.e. one which will cause Fur to bind DNA and repress transcription. No significant changes in chemical shift, linewidth or relaxation times of the aromatic protons of the 9 histidines were observed, suggesting that metal binding in Fur does not direcly involve histidine sidechains. This result is in agreement with studies of histidine mutants of Fur, which found no unambiguous evidence of the requirement of any of the histidines for metal activated DNA binding.

More recent experiments have utilized Co^{2+}, another activating metal, and its interactions with both intact Fur protein as well as the FurT88 domain. Figure 3.12 shows a number of 1D NMR spectra of Fur and FurT88. Panels A-C show 1D spectra of intact Fur protein in H_2O in the absence (lower spectrum) and presence (upper spectrum) of approximately one mole equivalent of Co^{2+}. Panel D shows the a region of the 1D spectra of FurT88 also in the absence and presence of Co^{2+}. The overall envelope of the spectrum of the Fur protein is unchanged by the addition of metal in panels A and C, except for some broadening of resonances, indicating that these metal-protein interactions do not result in denaturation of the native protein structure. This is also true for the FurT88 protein (data not shown). Panel B shows a vertically and horizontally expanded region of the Fur spectrum of panel A, with a number of new peaks appearing on addition of Co^{2+}.

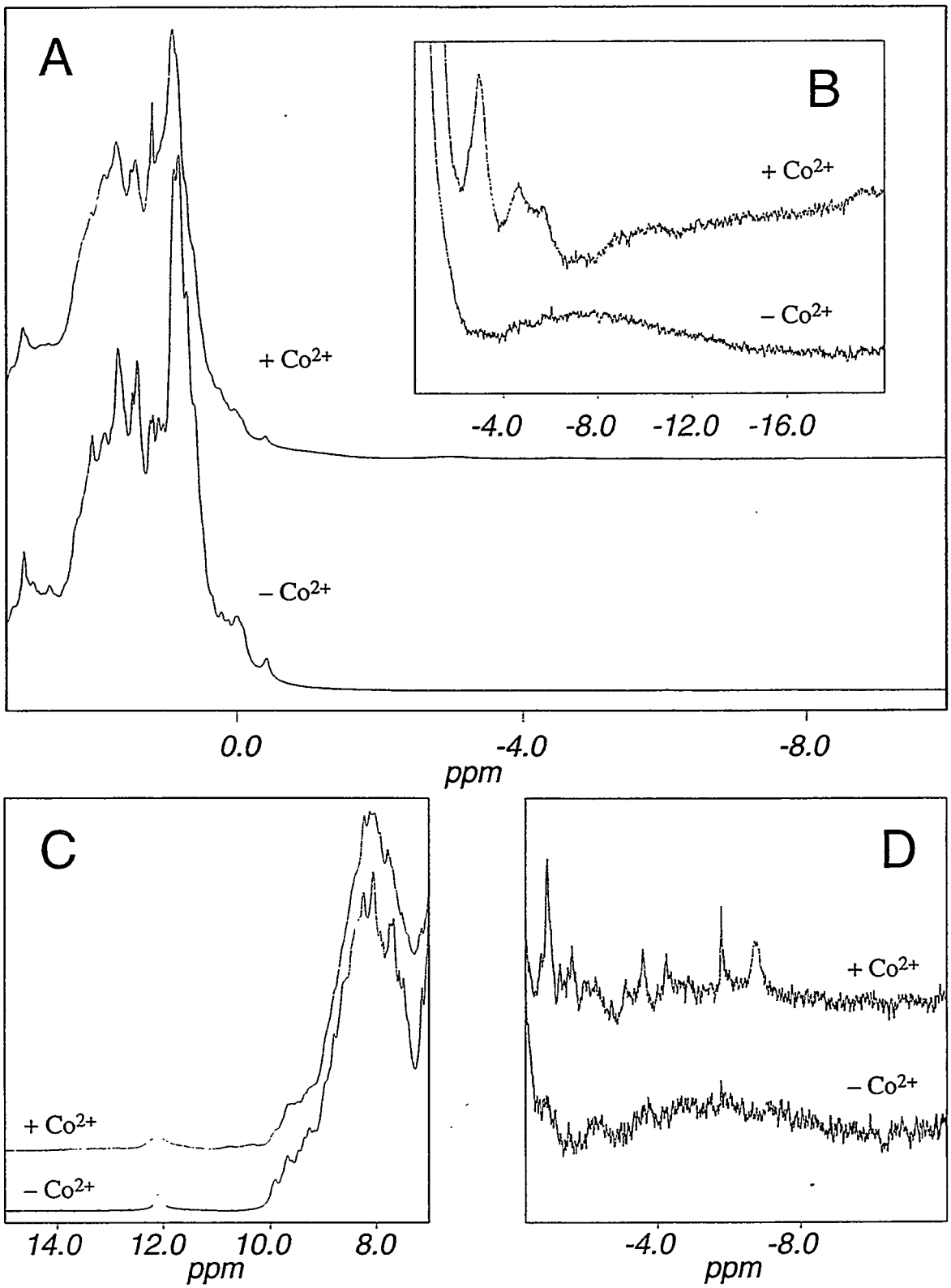


Figure 3.12: Effects of Co^{2+} on the NMR spectra of Fur and FurT88. Protein resonances which are shifted to high field values by the presence of metal are observed in both intact Fur protein (B) and the FurT88 metal-binding domain construct (D). See text for a complete description.

These are presumed to be resonances of the protein which have been shifted to high field by interactions with the metal which is bound specifically to the Fur protein. Similarly shifted signals are observed with FurT88, as shown in panel D. Additional new signals are observed in the regions from 10-30 ppm for both Fur and FurT88 (data not shown), providing another indication that specific metal-protein interactions are being detected.

Discussion

The NMR analysis of Fur protein structure and metal binding are currently in an intermediate phase. The results described in this chapter are not sufficient to allow the specific characterization of structural features, but are part of a continuing effort to understand the mechanism of metal-activated DNA binding by Fur. The results presented are, however, sufficient to point out discrepancies between the previous NMR studies of Fur (Saito et al., 1991) and the current view of the structure of the Fur metal-binding domain.

While the previous chemical shift assignments of Fur would not necessarily be expected to transfer directly to the assignments of FurT and FurT88 reported here, the similarity between the HMQC spectra of Fur and FurT (Figure 3.2) is suggestive that few major differences would be found. Comparison of NH and αH chemical shifts of residues for which assignments are available from both this work and the previous study reveals only one pair of residues with reasonable agreement, A115 and K116. Since the majority of assignments previously reported for Fur were sidechain protons, rather than backbone protons, and the assignments were performed on spectra collected in D_2O, it is not clear how accurately these resonances were assigned in the context of the sequence.

In addition to resonance assignments, a model for the three-dimensional structure of the Fur protein was proposed which consisted of only helical and turn structural elements and no regions of extended β-sheet structure. This seems to be a very unlikely characterization in light of the patterns of sequential NOEs and slowly exchanging amides

observed for FurT88, which indicate extended backbone conformation for residues whose amides are likely to be forming hydrogen bonds, as shown in Figure 3.11. The previous model also included no indication of dimerization determinants, which are now known to be contained in the metal-binding domain. It is anticipated that a more complete determination of the secondary structure of FurT88 will confirm the presence of both helix and sheet elements, rather than a completely helical structure, and provide a complete picture of the dimerization interface.

The other aspect of Fur which is of immediate interest is the identification of the residues which are directly involved in the liganding of activating metal ions. The preliminary results described above serve merely to show that NMR may be useful in directly studying the metal binding process, and to confirm that the FurT88 domain will be a good model for the interactions of Fur protein with activating metal, i.e. a satisfactory structural and functional mimic of the intact protein.

To reach these immediate goals a number of additional NMR experiments have been planned. The notable difference between the experiments described in this chapter and those to be performed is the switch to uniformly ^{13}C-^{15}N-labeled protein, and the use of a number of more recently developed techniques for obtaining resonance assignments and structural information. The approach to be taken will be similar to that described in Chapter 4 for the determination of the structure of the NTRC receiver domain. It is likely that additional selective labeling of amino acids will be used for assignment purposes and efforts to obtain specific metal binding site information. If, for instance, further evidence of involvement of cysteine sidechains in metal binding is found, it would be useful to incorporate labeled cysteine as a very specific probe of metal-protein interactions.

The function of Fur protein which has not been addressed in these NMR studies is its interaction with DNA. Fur-DNA interaction is certainly an interesting topic, but the studies of the isolated metal-binding domain of Fur has moved the focus away from that aspect of Fur function at this time. A fusion of the first 80 residues of Fur with the maltose

binding protein (MBP) was shown to be sufficient for specific, high affinity DNA recognition. Construction of an N-terminal fragment of Fur should be attempted to provide a smaller molecule for NMR studes, much as the FurT fragment is being used for analysis of the metal binding of Fur. This may also answer questions which have arisen in the studies of Fur and FurT about the nature of the folded structure of the N-terminal domain of Fur .

In the proteolysis studies of Fur (Coy & Neilands, 1991), it was observed that no large fragments from the N-terminus of Fur were detected, while large amounts of the C-terminal domain FurT could be recovered, suggesting that the N-terminal half of Fur might be partially unstructured in the absence of DNA. Inspection of the HMQC spectra of Fur and FurT in Figure 3.2 might support this hypothesis. Close comparison reveals that the peaks of the well-dispersed spectrum of FurT comprise a majority of the resolved peaks of the Fur spectrum, with many of the additional signals attributable to the N-terminal residues of Fur falling into a narrow region centered about the random coil ^{15}N and ^{1}H chemical shifts. This cannot be comfirmed without a more detailed analysis of the NMR spectra of Fur and a significant amount of sequential assignment information. One simple way to further evaluate this condition would be to examine the HMQC spectrum of Fur when bound to an appropriate DNA sequence, looking for a dramatic change in the dispersion of the large group of signals in the center of the spectrum of the free protein. At least some of these experiments will be attempted in the near future with a view toward opening other avenues of study of the relationship between structure and function of the ferric uptake regulation protein.

Chapter 4 The Three-Dimensional Solution Structure of the N-Terminal Receiver Domain of NTRC

Introduction

"Two-component" signal transduction pathways are remarkably common in eubacterial responses to changes in local environment (Parkinson & Kofoid, 1992) and have recently been identified in eukaryotic cells (Chang et al., 1993; Ota & Varshavsky, 1993; Maeda et al., 1994). Such two-component pathways are involved in a wide array of processes, such as nitrogen regulation, chemotaxis, and osmoregulation. The archetypal system consists of a histidine kinase and a response regulator or receiver domain. The histidine kinase autophosphorylates, often in response to an environmental stimulus. The phosphate incorporated is then transferred from the histidine in the autokinase to an aspartate residue in the receiver domain of the cognate second component. The phosphorylated receiver domain modulates the function of an attached output domain or, rarely, a separate protein.

In the case of nitrogen regulation, the phosphorylation of D54 in the amino (N)-terminal domain of the NTRC protein, the receiver domain in this system, is accomplished via a phosphotransfer reaction from the histidine kinase, NTRB (Ninfa & Magasanik, 1986; Keener & Kustu, 1988; Weiss & Magasanik, 1988; Sanders et al., 1992). The phosphorylation of D54 in the N-terminal domain of NTRC is required for activation of an ATPase in the central domain of the protein (Weiss et al., 1991; Austin & Dixon, 1992). This, in turn, allows NTRC to catalyze the formation of open complexes between the σ^{54}-holoenzyme form of RNA polymerase and promoters, resulting in the transcription of genes whose products participate in nitrogen metabolism. Phosphorylation of the N-terminal domain of NTRC is not required for its DNA-binding per se, but induces oligomerization of NTRC dimers at enhancers, which are composed of two binding sites for dimers; activation of the ATPase is thought to be mediated by this oligomerization

(Austin & Dixon, 1992; V. Weiss et al., 1992b; Porter, 1993; Porter et al., in press). Unlike phosphorylation, deletion of the receiver domain of NTRC fails to activate the ATPase of the central domain (Drummond et al., 1990; D. Weiss et al., 1992).

The three dimensional structure of CheY, a member of the receiver domain superfamily, has been solved by X-ray crystallography with and without Mg^{2+} bound (Stock et al., 1989; Volz & Matsumura, 1991; Stock et al., 1993; Bellsolell et al., 1994). Mutational analysis and sequence comparisons have indicated that the sidechains of residues D12, D13, D57, T87 and K107 form the active site of CheY, with D57 the site of phosphorylation (Lukat et al., 1991; Volz, 1993). Receiver domains themselves catalyze phosphate incorporation from their cognate autokinases and from low molecular weight donors such as carbamyl phosphate, acetyl phosphate, and phosphoramidate (Feng et al., 1992; Lukat et al., 1992). In addition, a number of them have been shown to have an autophosphatase activity (Hess et al., 1988; Keener & Kustu, 1988). CheY, in contrast to most two-component receiver domains, is a single domain protein that interacts with its target(s) in the switch complex of the flagellar motor to control the direction of flagellar rotation (Ravid et al., 1986). By contrast, NTRC, like most other members of the superfamily, contains the N-terminal receiver domain and its downstream target within the same protein. Differences in the structures of CheY and the N-terminal domain of NTRC may indicate regions important for the interaction of these receiver domains with their respective downstream targets.

The important regulatory role of receiver domains of two-component systems raises many questions about relationships between structure and function. Of primary interest are the structural rearrangements postulated to occur upon phosphorylation and communication to downstream targets via these rearrangements. Determination of the three-dimensional structure of the receiver domain of NTRC is a step toward obtaining detailed information on the mechanism of signal transduction from this domain to the central domain of the protein. In this chapter we report the NMR assignments and three-dimensional solution

structure of the N-terminal receiver domain of the NTRC protein from *Salmonella typhimurium* and discuss differences from CheY.

Materials and Methods

Expression, Purification and Enzymatic Activity of the NTRC Receiver Domain: The expression vector pJES592 (Klose et al., 1994), which includes a T7 promoter and a DNA fragment encoding the N-terminal domain of NTRC (residues 1-124), was transformed into *E. coli* BL21(DE3) cells carrying the pLysS plasmid (Studier et al., 1990). To obtain uniform labeling of protein samples, cells were grown on M9 minimal medium (Sambrook et al., 1989) at 37° C with $^{15}NH_4Cl$ and [$^{13}C_6$]-D-glucose as the sole sources of nitrogen and carbon, respectively. NTRC receiver domain selectively labeled with ^{15}N-leucine was grown similarly, but with the addition of the other 19 naturally occurring amino acids at a concentration of 100 mg/L, and the isotopically enriched amino acid at 150 mg/L. Production of the NTRC receiver domain was induced with the addition of 1 mM isopropyl β-D-thiogalactopyranoside after the cell density had reached 0.25 absorbance units at 595 nm. The cells were grown for 9-12 hours after induction, whereupon they were harvested by centrifugation. The cells were lysed by sonication in lysis buffer (100 mM KCl, 50 mM Tris-acetate, pH 8.2, 5% glycerol), and a crude extract was prepared by centrifugation at 20,000 rpm for 20 minutes in an SW28 rotor. The supernatant was diluted twofold and applied to a DEAE Sephadex-50 column. The column was washed with five column volumes of running buffer (50 mM NaCl, 10 mM phosphate, pH 6.8, 0.5 mM dithiothreitol) and eluted in a stepped gradient of increasing salt concentration (50-500 mM NaCl in 50 mM increments of 30 mL). The NTRC receiver domain elutes at 250-300 mM NaCl. The fractions containing NTRC receiver domain were concentrated using Centriprep-10 (Amicon) flow concentrators. Final HPLC purification was performed on a 5PW DEAE ion exchange column (Waters). Purity and identity of the protein were confirmed by mass spectroscopy and NMR.

The rate of phosphorylation of the receiver domain and its steady-state level of phosphorylation, which represents a balance between phosphate incorporation and release, were checked with 1 μM NTRC receiver domain [preparations before and after lyophilization (see below)] and 200 nM NTRB as described (Keener & Kustu, 1988). Activities of the receiver domain were similar to those of intact NTRC and a maltose-binding protein fusion to the receiver domain, verifying that preparations used for NMR spectroscopy had normal enzymatic activity.

Sample Preparation: Concentrated protein solution was flow dialyzed against 10mM phosphate buffer, pH 6.4, and lyophilized. Dry protein samples were dissolved in 0.5 mL D_2O or 10% D_2O/90% H_2O. The pH of NMR samples was adjusted to 6.4 with 1M HCl or NaOH. The concentrations of the uniformly ^{15}N-labeled and [^{15}N]Leu-labeled samples were 2 mM; the concentration of the uniformly ^{13}C,^{15}N-labeled sample was 3 mM. The concentrations of NMR samples are based on the weight of the lyophilized material after HPLC purification and dialysis against water.

NMR Experiments: NMR experiments were performed at 600 MHz on a Bruker AMX-600 spectrometer at 25° C. Chemical shift values were externally referenced to TSP (^1H and ^{13}C) (Driscoll et al., 1990) and liquid ammonia (^{15}N) (Live et al., 1984). Non-acquisition dimensions of all multidimensional experiments utilized the States-TPPI method for quadrature detection (Marion et al., 1989a). All data were processed with FELIX version 2.30β (Biosym), including linear prediction calculations. Shifted skewed sine-bell functions were used for apodization of the free induction decays.

^{15}N-edited 3D NOESY-HMQC (Kay et al., 1989; Marion et al., 1989b), and 3D TOCSY-HMQC (Driscoll et al., 1990), experiments were collected with spectral widths of 6944 Hz for the ^1H dimensions and 1861 Hz for the ^{15}N dimension. The ^1H carrier was placed on the H_2O resonance at 4.80 ppm, and the ^{15}N carrier set to 116.2 ppm. The NOESY mixing time was 100 ms, and the TOCSY spin-lock period was 80 ms. A total of 128 x 32 x 1024 complex points were collected in the t1, t2, and t3 dimensions,

respectively. Data were apodized in each dimension with a shifted, skewed sine-bell. A shift of 75° was used in each dimension, with a skew of 1.0, 0.8, and 0.5 in the t1, t2, and t3 dimensions, respectively. Data were zero-filled to yield a 512 x 64 x 512 real matrix upon Fourier transformation.

^{15}N-^{1}H 2D HSQC (Bodenhausen & Ruben, 1980; Marion et al., 1989a) experiments were collected with identical spectral parameters, but 256 complex points were acquired in the ^{15}N dimension to yield a high-resolution spectrum for assignment purposes. A 2D HMQC-J experiment was collected with similar parameters to the HSQC experiments, but with 498 complex points in order to obtain $J^{3}_{H\alpha\text{-}HN}$ values used to generate qualitative dihedral angle restraints (Kay & Bax, 1990). The ^{15}N-^{1}H HSQC experiment also provided amide exchange information from a sample dissolved in D_2O immediately prior to acquisition of a series of 2D experiments.

2D ^{1}H-^{13}C HSQC experiments were collected for both the aliphatic and aromatic resonances of the NTRC receiver domain, using a constant-time (CT) evolution period for ^{13}C equal to $1/J_{CC}$, producing a completely ^{13}C-decoupled spectrum in t1 (Vuister & Bax, 1992). The aromatic CT-HSQC was centered at 122.64 ppm ^{13}C, with a ^{13}C spectral width of 3968 Hz, and at 4.80 ppm ^{1}H, with a ^{1}H spectral width of 7246 Hz. The aliphatic CT-HSQC was centered at 43.16 ppm ^{13}C with a ^{13}C spectral width of 5000 Hz and at 4.80 ppm ^{1}H, with a ^{1}H spectral width of 7246 Hz. After normal processing, ^{1}H-^{13}C HSQC spectra were treated with an average noise measurement routine, ANI, and a t1 noise reduction routine, RT1, to attenuate the streaking of intense methyl and aromatic signals, which tended to obscure weaker peaks (Manoleras & Norton, 1992). Post-processing with ANI and TR1 used a threshold value, T, of 5 and a spread value, η, of 3.0 with 10 iterations.

A 3D ^{13}C HCCH-TOCSY experiment (Bax et al., 1990) was acquired with parameters identical to the ^{1}H dimension of the aliphatic 2D CT-HSQC, but in the ^{13}C dimension only 27 complex points were collected with a spectral width of 2809 Hz,

centered at 43.16 ppm. Extensive ^{13}C aliasing was used to maintain reasonably high resolution, despite the low digitization in that dimension. Methyl resonances which appear above 18.39 ppm in the ^{13}C dimension were aliased to values two spectral widths downfield. 128 complex points were collected in the t1 ^{1}H dimension, and zero-filled to yield a 256 x 64 x 512 real matrix.

A 4D ^{13}C HMQC-NOESY-HMQC experiment (Clore et al., 1991) with a 100 ms mixing time was used to generate distance restraints between carbon-bound protons. Eight complex points in each of the two ^{13}C dimensions (t1 and t3), 56 points in the t2 dimension, and 256 points in the t4 dimension were acquired over 76 hours. Both ^{1}H dimensions were centered at 4.13 ppm with spectral widths of 5319 Hz. The ^{13}C dimensions were centered at 43.16 ppm with spectral widths of 2809 Hz. In processing the data, both ^{1}H dimensions were processed normally, followed by Fourier transformation of the t3 ^{13}C dimension without apodization. The t1 ^{13}C dimension was then extended to 12 complex points with linear prediction, apodized and Fourier transformed, followed by inverse transformation of the t3 dimension, linear prediction, apodization and Fourier transformation. The time domain data were zero-filled to produce a 16 x 128 x 16 x 256 real matrix.

A 3D CBCA(CO)NH (Grzesiek & Bax, 1992a) experiment was collected with ^{1}H and ^{15}N parameters identical to the 3D ^{15}N experiments described above. The ^{13}C dimension was centered at 43.16 ppm with a spectral width of 8446 Hz. 50 complex points were collected in the ^{13}C dimension and linear predicted to 75 points. Time-domain data were zero-filled to yield a 256 x 64 x 512 real matrix.

Structure Calculations: Structure calculations were performed using the program X-PLOR 3.1 (Brünger, 1992). A standard protocol for embedding, annealing and optimizing the coordinates was used (Nilges et al., 1988; Brünger, 1992). The *dg_sub_embed* routine was used to generate starting structures with substructures embedded from the unsmoothed bounds matrix. Molecular dynamics and simulated annealing were performed with the

dgsa routine, and the *refine* routine was used for final optimization of structures. Distance restraints were generated from the cross peaks of both the ^{15}N 3D NOESY-HMQC and the ^{13}C HCCH-NOESY, and classified as strong (1.8-2.7 Å), medium (1.8-3.5 Å) and weak (1.8-5 Å) (Williamson et al., 1985; Clore et al., 1986). Corrections were added to the upper bounds of restraints involving pseudoatoms for methylene and methyl groups, as well as Tyr and Phe ring protons (Wüthrich et al., 1983; Clore et al., 1986).

Dihedral angle restraints for selected ϕ angles were included on the basis of $J^3_{H\alpha\text{-}HN}$ coupling constants measured as splittings in the t1 dimension of the HMQC-J experiment. For large values of $J^3_{H\alpha\text{-}HN}$ (> 8 Hz) ϕ angles were constrained to -120° ± 40°, and for small values (< 6 Hz) ϕ angles were constrained to –60° ± 30°, if the residue was known to fall in a helical region of the protein, since other values of ϕ may give rise to small $J^3_{H\alpha\text{-}HN}$ values.

Structure refinement was performed in an iterative fashion, using each successive level of refinement to screen potential distance constraints on the basis of proximity in the structure. Hydrogen bonds were included as pairs of constraints between NH and N atoms to the corresponding carbonyl O atom, but only when NOE patterns indicated unambiguous donor-acceptor pairs and if the NH was observed in ^{15}N-^{1}H HSQC spectra collected at least 1 hour after dissolving the sample in D_2O. Hydrogen bonds were defined by restraints defining the O-N distance to be between 2.8 Å and 3.3 Å and the O-H distance to be between 1.8 Å and 2.3 Å.

Results

Resonance Assignments: Figure 4.1 shows the ^{15}N-^{1}H HSQC spectrum of the NTRC receiver domain, with assigned peaks labeled by residue and number. The single set of resonances with narrow linewidths is consistent with the NTRC receiver domain being monomeric in solution at the concentrations used. Sequence-specific assignment of

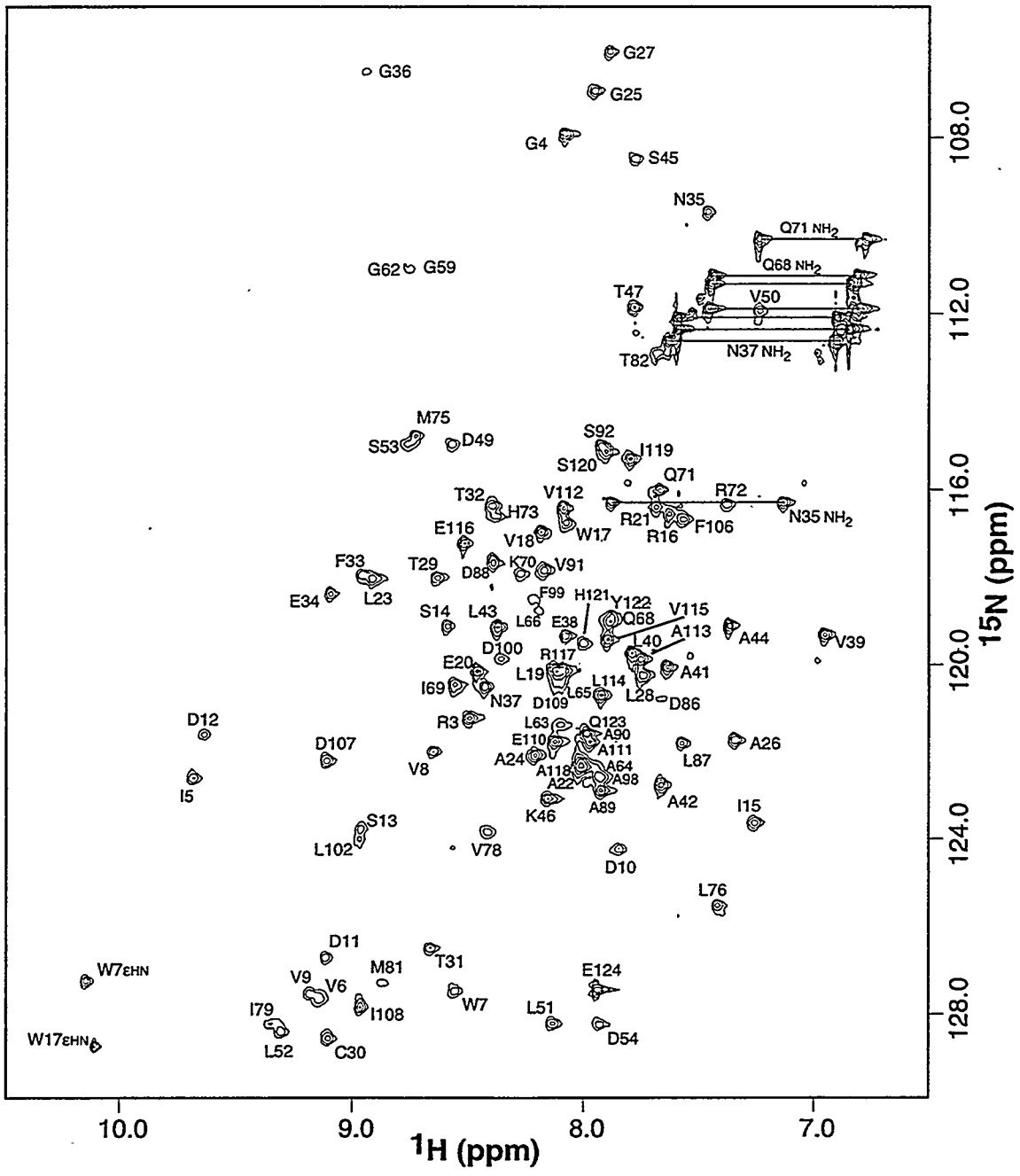


Figure 4.1: ^{1}H-^{15}N HSQC spectrum of uniformly ^{15}N-labeled NTRC receiver domain. Cross-peak assignments are indicated with the one-letter amino acid code and residue number.

^{1}H and ^{15}N resonances was completed using primarily the ^{15}N 3D NOESY-HMQC and ^{15}N 3D TOCSY-HMQC data collected from uniformly ^{15}N-labeled protein. The separation of spin-systems by their amide ^{15}N chemical shift allows the straightforward identification of sequential αH-NH, NH-NH, and βH-NH NOEs for assignment of residues in a traditional manner (Wüthrich, 1986). Figure 4.2 contains selected strips from the ^{15}N 3D NOESY-HMQC and ^{15}N 3D TOCSY-HMQC experiments, illustrating the method by which ^{15}N-directed sequential assignments were made. The wide αH chemical shift dispersion which simplified much of the sequential assignment is also visible in Figure 4.2, with crosspeaks from W7 Hα (6.55 ppm) and V39 Hα (2.84 ppm) indicated. Additional assignments were aided by comparison of the HSQC spectrum of [^{15}N]Leu-labeled protein with the uniformly labeled spectrum to unambiguously determine the residue type for those unassigned leucine residues. A number of ^{15}N-^{1}H correlations are not observed, possibly due to high exchange rates at near-neutral pH. A total of 85% of all backbone amide ^{15}N and ^{1}H resonances were assigned from these data, and a significant portion of the αH and sidechain protons were assigned from the ^{15}N 3D TOCSY-HMQC spectrum.

One plane of the ^{13}C 3D HCCH-TOCSY experiment is shown in Figure 4.3. In order to assign the aliphatic ^{13}C resonances in the protein and complete the sidechain ^{1}H assignments, ^{13}C-$^{1}H\alpha$ peaks from the CT-HSQC experiment were correlated with spin systems in the ^{13}C 3D HCCH-TOCSY and then matched with Hα and sidechain assignments from the ^{15}N 3D TOCSY-HMQC spectrum. Analysis of the 3D CBCA(CO)NH experiment correlated $^{13}C\alpha$ and $^{13}C\beta$ resonances with the ^{15}N and ^{1}H resonances of the following sequential residue, providing additional $^{13}C\alpha$ and $^{13}C\beta$ assignments for prolines and some residues whose ^{15}N-^{1}H correlations were not observed. These additional assignments were matched with unassigned spin systems remaining in the ^{13}C 3D HCCH-TOCSY data. Chemical shift values for most resonances were tabulated

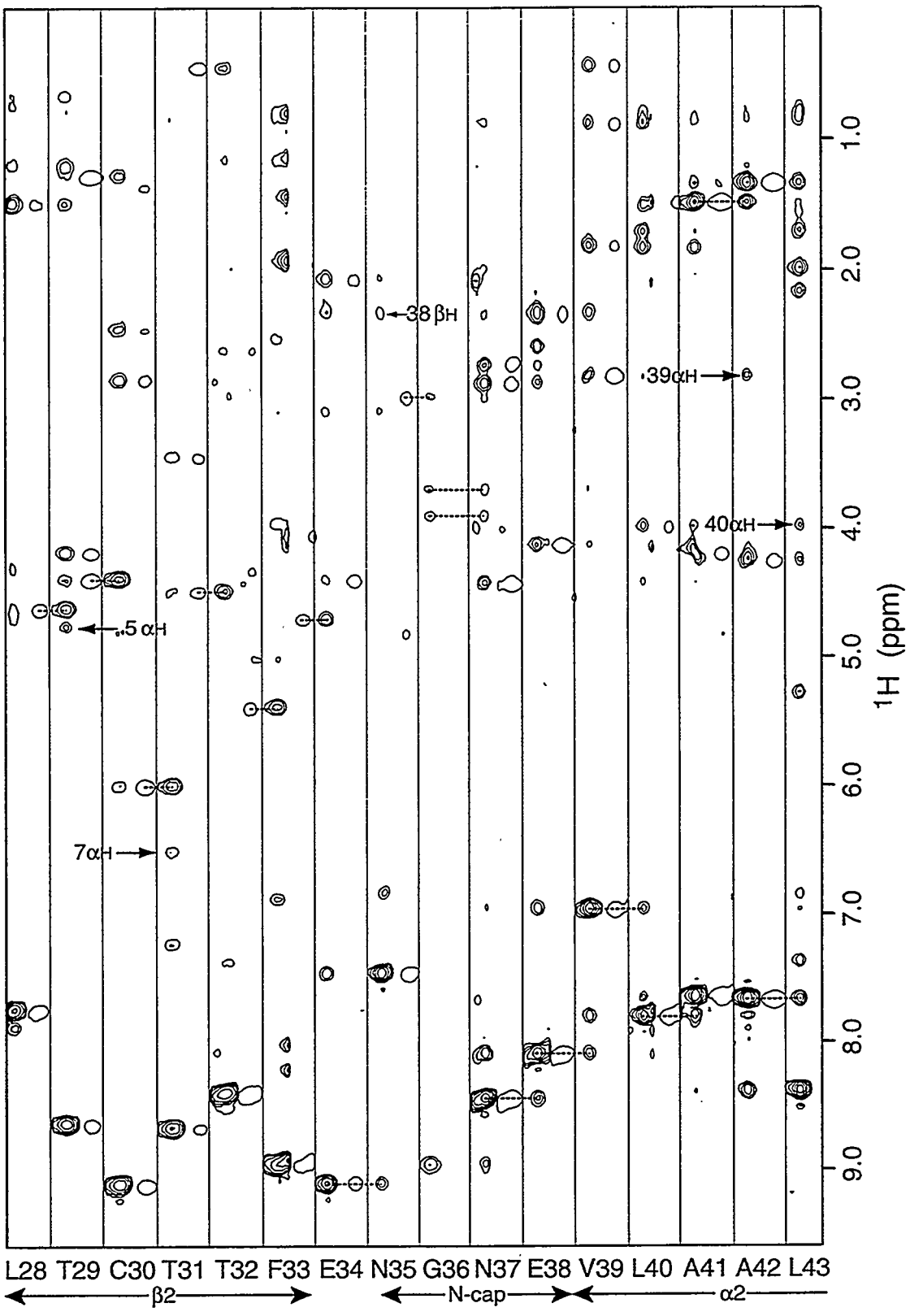

38βH
39αH
40αH
5αH
7αH
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
^{1}H (ppm)
L28 T29 C30 T31 T32 F33 E34 N35 G36 N37 E38 V39 L40 A41 A42 L43
β2
N-cap
α2

Figure 4.2: Strip plot of selected regions of the 3D ^{15}N NOESY-HMQC and TOCSY-HMQC spectra of NTRC receiver domain illustrating the sequential resonance assignments for residues 28-43. Each pair of vertical strips is taken from the 3D ^{1}H-^{1}H plane whose ^{15}N value is the amide nitrogen chemical shift of the amino acid listed on the horizontal axis. Strips are centered in F3 at the amide proton chemical shift of the assigned residue. NOESY and TOCSY strips are alternated, with NOESY peaks displayed with multiple contours and TOCSY peaks drawn with a single contour. Sequential connectivities are indicated with dashed lines between TOCSY and NOESY peaks of neighboring residues. Residues 28-34 and 35-36 are connected by αH-NH NOEs; residues 34-35, 37-41 and 42-43 are connected by NH-NH NOEs; residues 41-42 are connected by a βH-NH NOE, due to the degeneracy of their amide protons. Labeled arrows denote crosspeaks due to medium and long range NOEs between the specified proton and the amide proton of the residue indicated on the horizontal axis. Crosspeaks from residues 5 to 29 and 7 to 31 are due to cross-strand NOE's in the parallel β-sheet structure. A βH-NH NOE from E38 to N35 is indicative of an N-capping interaction in helix 2. Helical αH-NH(i, i+3) NOEs from 39 to 42 and 40 to 43 are also shown.

from peaks in the high-resolution HSQC experiments, with the remainder measured in the ^{13}C 3D HCCH-TOCSY or ^{13}C 4D NOESY data.

The 4D ^{13}C HMQC-NOESY-HMQC (4D NOESY) experiment was used to identify connections between proline residues and the preceding residues, as αH-αH or αH-δH sequential NOEs are observed for 5 of the 6 proline residues of the NTRC receiver domain. The conserved *cis*-peptide bond between K104 and P105 was confirmed in this manner, in agreement with the crystal structures of CheY. The exception is P48, for which complete resonance assignments were not obtained. Its $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts were determined by correlation with the NH of D49 in the CBCA(CO)NH experiment, but it appears to be highly degenerate with other spin systems in the ^{13}C 3D HCCH-TOCSY. No NOEs to P48 could be identified which would confirm either the *cis*- or *trans*- form of the T47-P48 peptide bond; however, the value of the $^{13}C\beta$ chemical shift has been shown to be a reliable indicator of proline peptide bond isomerization (Scanlon & Norton, 1994). The value observed for the $^{13}C\beta$ of P48 is 30.35 ppm, in agreement with the value of 30.6 ppm reported for proline in a *trans*- configuration (Dorman & Bovey, 1973; Wüthrich, 1976).

The 4D NOESY was also helpful in confirming assignments of aromatic ^{13}C and ^{1}H resonances, by identifying strong crosspeaks between aromatic protons and assigned βH resonances of aromatic residues. Tentative assignments from the 4D NOESY were compared with peaks in the aromatic ^{13}C CT-HSQC spectrum to confirm that ^{13}C values were in the appropriate ranges for the various types of aromatic sidechains. Table 4.1 contains ^{1}H, ^{15}N and ^{13}C chemical shift assignments for the receiver domain of NTRC at pH 6.4 and 25° C. In total, more than 90% of all ^{1}H, ^{15}N and ^{13}C resonances of NTRC receiver domain were assigned sequence-specifically, with unobserved backbone amides and the corresponding sidechains being the majority of missing assignments.

Identification of Secondary Structure: The location of secondary structure elements was determined initially by analysis of NOE patterns in the ^{15}N NOESY data. Figure 4.4

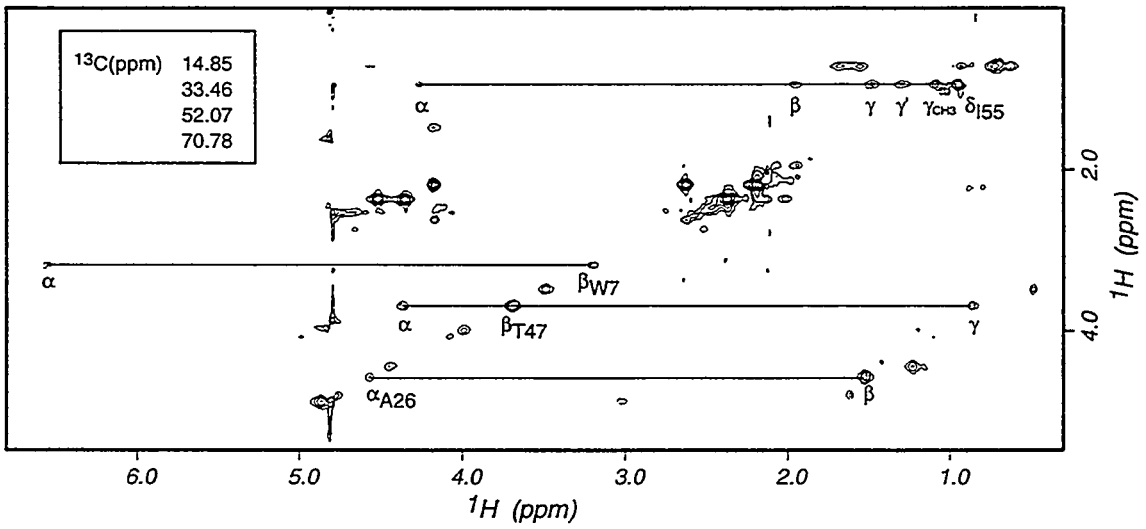


Figure 4.3: 2D ^{1}H-^{1}H plane of a 3D ^{13}C HCCH-TOCSY. Spin systems are labelled with residue assignment at the diagonal peak, with crosspeaks labeled by proton type. Due to extensive aliasing in the F2 dimension, the ^{13}Cδ of I55 at 14.85 ppm, the ^{13}Cβ of W7 at 33.46 ppm and the ^{13}Cα of A26 at 52.07 ppm appear in the same plane as ^{13}Cα of T47 at 70.78 ppm. Horizontal lines connect crosspeaks from all protons in the spin system, illustrating the usefulness of this experiment in identifying amino acid type by the pattern of peaks observed.

Table 4.1 ^{1}H, ^{15}N, and ^{13}C resonance assignments of the NTRC Receiver Domain at 25° C, pH 6.4

Ala	N	HN	CA	HA	CB	HB*
A22	122.44	8.00	54.72	4.22	19.36	1.47
A24	122.09	8.21	52.61	4.78	19.26	1.61
A26	121.74	7.34	51.60	4.56	18.87	1.52
A41	120.06	7.64	54.92	4.19	18.11	1.54
A42	122.76	7.66	55.13	4.27	19.55	1.37
A44	119.11	7.36	54.36	4.31	18.59	1.62
A64	122.03	7.92	52.27	4.42	19.10	1.21
A83			55.14	4.74	17.76	1.58
A89	122.95	7.91	54.60	4.18	18.46	1.48
A90	121.83	7.97	54.72	4.01	18.54	1.40
A93	119.81	7.54	55.17	4.20	18.50	1.48
A98	122.66	7.93	54.20	4.17	18.52	1.40
A111	122.29	7.97	55.40	3.84	18.40	1.07
A113	119.89	7.75	55.12	4.21	17.92	1.51
A118	122.40	8.03	55.00	4.17	18.43	1.50

Arg	N	HN	CA	HA	CB	HB1	HB2	CG	HG1	HG2	CD	HD1	HD2
R 3	121.23	8.50	55.87	4.46	31.51	1.85	1.76	27.29	1.69	1.63	43.27	3.18	3.18
R 16	116.56	7.62	61.73	3.80	30.48	2.16	2.17				43.73	2.99	2.97
R 21	115.38	7.68	57.11	4.05	31.40	2.04	1.92		1.60	1.60	42.00	3.12	3.07
R 56					30.55	1.96	1.85	27.32	1.75	1.62	43.20	3.32	3.27
R 72	116.34	7.37	57.18	4.12	31.96	1.70	1.61				43.35	3.17	3.11
R 117	120.15	8.09	59.46	4.15	30.48	2.17	2.15	28.29	1.98	1.75	44.04	3.33	3.32

Asn	N	HN	CA	HA	CB	HB1	HB2	HND1	HND2	ND
N 35	109.70	7.46	52.17	4.86	41.40	3.00	3.00	7.13	7.88	116.28
N 37	120.51	8.43	56.53	4.45	37.44	2.90	2.76	7.63	6.90	112.63

Asp	N	HN	CA	HA	CB	HB1	HB2
D 10	124.25	7.84	54.78	4.65	44.36	2.57	2.43
D 11	126.74	9.11	55.78	4.62	42.02	2.95	2.74
D 12	121.61	9.64	53.55	4.78	41.32	3.12	2.65
D 49	114.97	8.57	56.50	4.53	42.73	2.67	2.32
D 54	128.26	7.93	54.52	5.40	41.89	2.96	2.96
D 61							
D 86	120.78	7.67	54.13	4.99	42.09	3.02	2.70
D 88	117.66	8.39	56.91	4.36	40.27	2.65	2.65
D 100	119.88	8.36		4.13		2.49	2.24
D 107	122.21	9.11	53.30	4.91	42.71	2.87	2.73
D 109	120.42	8.11	57.64	4.47	40.00	2.81	2.67

Table 4.1 (cont.): ^{1}H, ^{15}N and ^{13}C resonance assignments for the NTRC receiver domain, at 25° C, pH 6.4

Cys	N	HN	CA	HA	CB	HB1	HB2	HSG
C 30	128.57	9.11	56.52	6.03	29.59	2.87	2.48	1.36

Gln	N	HN	CA	HA	CB	HB1	HB2	CG	HG1	HG2	HNE1	HNE2	NE
Q 2			55.57	4.51	30.06	2.12	2.01	33.82	2.36	2.36			
Q 68	119.01	7.88	59.28	4.15	28.65	2.31	2.21	34.08	2.56	2.42	7.43	6.81	111.14
Q 71	116.02	7.67	58.35	4.06	28.75	2.23	2.17	34.11	2.56	2.42	7.23	6.77	110.31
Q 95			58.21	4.12	28.13	2.17	2.17	34.02	2.45	2.45			
Q 96			55.17	4.59	31.31	2.05	2.05						
Q123	121.56	7.99	55.62	4.35	29.90	2.14	2.00	33.79	2.36	2.36			

Glu	N	HN	CA	HA	CB	HB1	HB2	CG	HG1	HG2
E 20	120.15	8.46	60.74	4.03	29.19	2.12	2.12	36.41	2.27	2.22
E 34	118.38	9.10	56.79	4.44	31.55	2.12	2.12		2.26	2.37
E 38	119.34	8.08	58.82	4.15	30.71	2.36	2.36	36.47	2.63	2.41
E 110	121.78	8.12	58.77	4.13	29.41	2.24	2.19	36.56	2.44	2.31
E 116	117.24	8.52	59.89	3.87	28.70	2.16	2.04	35.92	2.55	2.16
E 124	127.46	7.93	58.10	4.11	31.05	2.08	1.94	36.73	2.28	2.28

Gly	N	HN	CA	HA1	HA2
G 4	107.92	8.08	45.38	4.02	3.95
G 25	106.95	7.95	46.52	4.02	3.91
G 27	106.06	7.88	45.29	4.36	3.80
G 36	106.51	8.94	47.05	3.93	3.73
G 59	110.99	8.76	45.53	4.12	3.87
G 62	110.99	8.75	47.15	4.17	3.88
G 97			45.38	4.03	4.03

His	N	HN	CA	HA	CB	HB1	HB2	CD2	HD2	CE1	HE1
H 73	116.61	8.37	53.19	5.04	30.18	3.02	3.02	121.60	6.84	139.79	7.95
H 84			59.39	4.38	29.37	3.40	3.34	120.54	7.31	138.30	8.31
H 121	119.50	7.80	56.60	4.56	28.58	3.28	3.28	119.97	6.99	137.62	8.46

Ile	N	HN	CA	HA	CB	HB1	CG1	HG11	HG12	CG2	HG2	CD	HD*
I 5	122.59	9.68	60.74	4.78	39.70	1.66	27.49	1.68	1.08	18.50	1.07	13.38	0.95
I 15	123.65	7.25	61.35	4.02	36.54	2.28	27.07	1.51	1.46	18.54	1.05	9.47	0.71
I 55			61.70	4.26	39.29	1.95	28.29	1.49	1.28	18.54	1.07	14.53	0.94
I 69	120.47	8.56	66.05	3.61	37.81	2.04	29.41	1.88	0.97	18.05	0.86	13.82	0.72
I 79	128.25	9.34	59.60	4.58	40.26	1.69				18.46	0.62.	14.65	0.72
I 80													
I108	127.85	8.97	62.82	4.04	37.92	2.04	28.74	1.52	1.49	18.48	1.03	13.96	1.02
I119	115.32	7.80	63.77	3.79	38.09	1.91	28.85	1.62	1.01	17.27	0.85	14.14	0.68

Table 4.1 (cont.): ^{1}H, ^{15}N and ^{13}C resonance assignments for the NTRC receiver domain, at 25° C, pH 6.4

Leu	N	HN	CA	HA	CB	HB1	HB2	CG	HG	CD1	HD1	CD2	HD2
L 19	120.11	8.13	57.81	4.17	42.01	2.08	1.25	28.43	1.97	25.83	0.91	22.37	0.66
L 23	118.02	8.92	57.98	4.11	39.92	1.98	1.98			21.40	1.21	26.85	0.85
L 28	120.26	7.75	53.84	4.65	42.59	1.52	1.25	27.49	1.50	26.18	0.78	23.76	0.70
L 40	119.75	7.79	57.98	4.02	41.31	1.87	1.56	27.04	1.76	24.17	0.92	22.96	0.82
L 43	119.88	8.37	55.55	5.32	42.85	2.04	1.72	27.50	2.22	26.90	0.88	24.85	0.81
L 51	128.24	8.13	53.40	5.40	46.64	1.84	1.15	28.31	1.52	25.99	1.03	26.85	0.85
L 52	128.41	9.31	54.53	5.59	44.68	1.79	1.56	29.42	1.59	25.83	0.80	26.51	0.71
L 63	120.57	8.09	57.00	4.27	39.28	1.75	1.75	28.80	1.50			24.14	0.92
L 65	119.90	8.10	59.04	4.14	41.36	1.77	1.59	27.11	1.75	25.40	0.89		
L 66	118.81	8.20	58.50	3.91						24.51	0.81	24.96	0.74
L 76	125.57	7.41	53.19	4.39	43.84	1.99	1.33	26.70	1.33	24.71	0.47	25.87	0.48
L 87	121.80	7.57	58.05	4.01	42.47	1.70	1.62	27.25	1.63	25.40	0.89	24.52	0.84
L 102	124.01	8.97	50.59	4.79	44.97	1.55	1.36	26.63	0.87	22.73	0.67	25.24	0.15
L 114	120.71	7.92	57.97	4.17	41.92	1.90	1.76			26.94	0.91	24.50	1.05

Lys	N	HN	CA	HA	CB	HB1	HB2	CG
K 46	123.08	8.15	55.67	4.67	36.50	2.08	1.98	24.77
K 67			59.98	4.07	32.36	2.01	1.98	25.75
K 70	117.91	8.28	58.56	3.95	29.15			26.44
K 104			53.31	4.59				24.74
	HG1	HG2	CD	HD1	HD2	CE	HE1	HE2
K 46	1.59	1.54	29.22	1.88	1.88	42.01	3.11	3.11
K 67	1.72	1.47	29.45	1.73	1.73	41.88	2.97	2.97
K 70	1.34	1.33	30.04	1.68	1.55	43.06	2.78	2.78
K 104	1.58	1.53				42.06	3.09	3.02

Met	N	HN	CA	HA	CB	HB1	HB2	CG	HG1	HG2	CE	HE*
M 1												
M 57			53.55	4.92	34.05	2.13	2.00	32.55	2.65	2.49	17.30	2.01
M 60												
M 75	114.81	8.73	54.21	4.65	31.44	2.31	2.06	32.76	2.74	2.51	18.30	2.08
M 81	127.31	8.87	53.67	5.61	35.66	2.08	1.92	32.16	2.57	2.40	18.10	2.11

Phe	N	HN	CA	HA	CB	HB1	HB2	CD*	HD*	CE*	HE*	CZ	HZ
F 33	117.99	8.96	56.62	4.73	43.59	3.15	2.58	131.06	6.93	130.17	6.36	129.05	5.39
F 99	118.51	8.21	59.64	4.55	38.71	3.27	3.07	132.87	7.12	136.49	6.69		
F 106	116.66	7.56	53.47	5.25	40.74	3.52	3.15	133.44	7.06	131.00	7.20		

Pro	CA	HA	CB	HB1	HB2	CG	HG1	HG2	CD	HD1	HD2
P 48	61.05		30.35								
P 58	63.71	4.46	31.98	2.36	2.03	27.74	2.24	2.11	50.86	3.95	3.84
P 74	65.35	4.59	32.32	2.50	2.02				50.11	3.66	3.32
P 77	62.97	4.52	31.82			28.14	2.21	1.77	51.53	4.16	3.97
P 103	62.10	4.96	32.11	2.14	1.88	27.75	2.24	2.24	50.66	3.80	3.61
P 105	62.57	4.39	34.70	2.25	1.96	27.55	2.04	2.04	50.35	3.72	3.52

Table 4.1 (cont.): ^{1}H, ^{15}N and ^{13}C resonance assignments for the NTRC receiver domain, at 25° C, pH 6.4

Ser	N	HN	CA	HA	CB	HB1	HB2
S 13	123.80	8.96	61.28	3.83	62.68	3.94	3.85
S 14	119.11	8.59	61.73	4.43	62.60	4.06	4.03
S 45	108.49	7.78	58.71	4.82	65.49	4.03	3.94
S 53	115.00	8.76	56.08	5.69	66.58	3.40	3.33
S 85			59.83	4.30	61.21	4.00	3.93
S 92	115.04	7.92	60.74	4.32	63.03	4.00	3.94
S 120	115.28	7.90	59.84	4.34	63.47	3.97	3.95

Thr	N	HN	CA	HA	CB	HB	CG	HG*
T 29	118.02	8.63	62.46	4.42	69.55	4.21	21.64	1.31
T 31	126.52	8.66	62.02	4.51	70.66	3.48	22.57	0.48
T 32	116.58	8.38	59.24	5.42	71.04	3.99	21.92	1.19
T 47	111.86	7.78	57.98	4.35	70.70	3.69	21.98	0.85
T 82	112.89	7.68	59.65	4.99	70.21	4.07	19.78	1.09

Trp	N	HN	CA	HA	CB	HB1	HB2	CD1	HD1
W 7	127.49	8.56	53.03	6.55	33.21	3.21	3.18	124.18	7.24
W 17	116.76	8.08	60.52	4.55	29.32	3.52	3.43	127.91	7.32
	NE1	HE1	CZ2	HZ2	CH2	HH2	HE3	CZ3	HZ3
W 7	127.26	10.15	114.31	7.20	114.58	7.27	6.95	118.12	6.85
W 17	128.76	10.11	115.06	7.47	124.46	7.24	7.66	121.83	7.16

Tyr	N	HN	CA	HA	CB	HB1	HB2	CD*	HD*	CE*	HE*
Y 94											
Y 101			56.11	5.53	41.73	2.81	2.74	132.60	6.86	117.80	6.67
Y 122	119.01	7.89	58.67	4.50	38.62	3.15	3.00	134.29	7.25	118.14	6.85

Val	N	HN	CA	HA	CB	HB	CG1	HG1*	CG2	HG2*
V 6	127.65	9.14	59.66	4.92	34.77	1.92	21.67	0.90	21.64	0.86
V 8	121.01	8.65	60.76	5.17	35.99	1.87	23.54	0.90	23.54	0.90
V 9	127.56	9.16	59.72	5.05	33.16	2.23	19.86	0.89	21.42	0.80
V 18	116.95	8.18	65.89	3.75	31.84	2.36	22.45	1.26	21.39	1.24
V 39	119.31	6.95	65.07	2.84	31.23	1.86	21.31	0.92	23.53	0.47
V 50	111.95	7.23	60.80	4.38	34.86	1.98	21.40	0.86	21.93	0.82
V 78	123.86	8.41	60.37	5.11	34.93	2.03	21.61	1.00	21.57	0.83
V 91	117.84	8.17	65.33	3.70	31.80	2.04	21.14	0.90	21.62	0.73
V 112	116.42	8.08	67.31	3.24	31.68	2.13	23.13	0.97	21.40	0.86
V 115	119.43	7.90	67.34	3.28	31.04	2.22	24.45	0.95	22.96	0.82

presents a summary of sequential and medium-range NOE information, as well as $J^3_{\alpha H\text{-}NH}$ coupling constant data and amide exchange information. Patterns of sequential βH-NH and NH-NH NOEs as well αH-NH(i, i+3) and αH-βH(i, i+3) NOEs indicate the presence of helices in the regions from residues 14-27, 36-44, 65-73, 85-95, and 108-121. Some αH-NH(i, i+3) NOEs for the start of helix 2 are highlighted in Figure 4.2. Strong sequential αH-NH NOEs indicate regions of extended conformation for the backbone of residues 2-11, 28-34, 50-54, 77-82 and 101-103. The large number of slowly exchanging amide protons in these regions suggests the presence of β-sheet structure.

Figure 4.5 shows the clear correlation of $^{13}C\alpha$ and $^{13}C\beta$ secondary shifts with the secondary structure elements of the NTRC receiver domain, especially when the differences between $^{13}C\alpha$ and $^{13}C\beta$ secondary shifts are examined in panel C. This agrees with the observed relationship between $^{13}C\alpha$ and $^{13}C\beta$ secondary shifts and location within an α-helix or β-sheet (Spera & Bax, 1991). The five regions of helical secondary shifts coincide with the pattern of helical NOEs shown in Figure 4.4, and the β-strands are in regions in which the trend of secondary shifts is reversed.

Helices are often bounded by initiation or termination signal sequences (Presta & Rose, 1988). The convention for specifying positions in and around helices is (...N'', N', N-cap, N1, N2 ... C2, C1, C-cap, C', C''...), with N-cap and C-cap denoting the first and last residues of the helix, respectively. Patterns of $^{13}C\alpha$ and $^{13}C\beta$ secondary shifts at the beginning of α-helices have been shown to indicate the presence of N-terminal helix capping interactions (Gronenborn & Clore, 1994). Close examination of the secondary shifts of the NTRC receiver domain indicate the possibility of capping interactions for helices 2 and 5, due to the characteristic downfield shift of the $^{13}C\beta$ resonances of the N-cap residues, N35 and D107, coupled with a slight shift upfield for the $^{13}C\alpha$ resonances of the same residues. The following three residues, at the N1-N3 positions of each helix, display downfield $^{13}C\alpha$ secondary shifts, in agreement with the reported pattern.

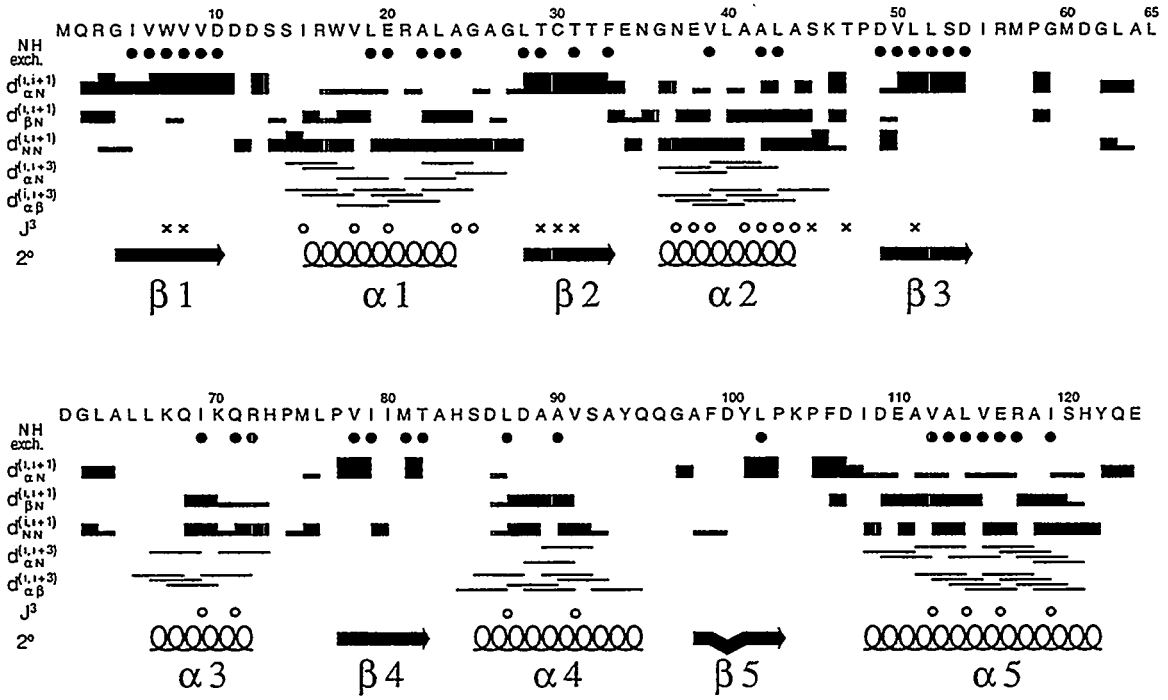


Figure 4.4: Summary of sequential and medium-range NOEs, slowly exchanging amide protons, and $J^3_{\alpha H\text{-}NH}$ values for NTRC receiver domain. Relative intensity of sequential NOEs is indicated by height of connecting box. Horizontal lines represent helical medium-range NOEs. Filled circles reflect amide protons which are observed in spectra collected at pH 6.4, 25° C, at least one hour after dissolving the sample in D_2O. Small (<6 Hz) and large (>9 Hz) values of $J^3_{\alpha H\text{-}NH}$ are represented by o and x, respectively.

NOEs characteristic of N-terminal capping interactions have also been identified (Lyu et al., 1993; Zhou et al., 1994b). NOEs are observed between Hβ resonances of the N3 residues, E38 and E110, and the backbone NH and sidechain resonances of the capping residues, N35 and D107, respectively. Interestingly, the N3 residues for both helices, E38 and E110, are potential "capping box" residues, having the ability to accept a reciprocal hydrogen bond from the backbone NH of the N-cap residue (Harper & Rose, 1993). In contrast to helices 2 and 5, neither the $^{13}C\alpha$ and $^{13}C\beta$ secondary shifts nor the NOE data for the N-terminal residues of helix 1 provide evidence of N-capping, and helix capping interactions could not be identified for the N-terminal residues of helices 3 and 4, due to the lack of NOEs defining the initiation points of those helices.

One instance of C-terminal helix capping is observed for residues 23-28, as evidenced by the slow exchange of the backbone NH of L28, and a pattern of NOEs which results in the positioning of L28 NH within 3.5 Å of L23 CO, and G27 NH within 2.0 Å of A24 CO. Recently, the termination of α-helices involving glycine residues has been classified into two major motifs (Aurora et al., 1994). The sequence of residues for the C-cap of helix 1 follows the proposed rules for the Schellman motif, which require a glycine at the C' position (G27) and apolar residues with hydrophobic contacts at the C3 and C" positions (L23 and L28). This arrangement produces a 6-1 (L28-L23), 5-2 (G27-A24) hydrogen bonding arrangement resulting in energetically favorable helix termination (Schellman, 1980).

Figure 4.6 displays the arrangement of the parallel β-sheet of NTRC receiver domain, as indicated by long-range NOEs, and the pattern of solvent-protected backbone amide protons. The five-stranded sheet has regular patterns of cross-strand connectivities, including αH-NH, αH-αH and NH-NH NOEs. Examples of cross-strand αH-NH NOEs can be seen in Figure 4.2. The N-terminal residue of strand 3, D49, plays an unusual role in the β-sheet. The NH of D49, while protected from solvent exchange, has NOEs to I5 NH and V50 NH. The I5 NH, which is also protected, has NOEs to the βH resonances of

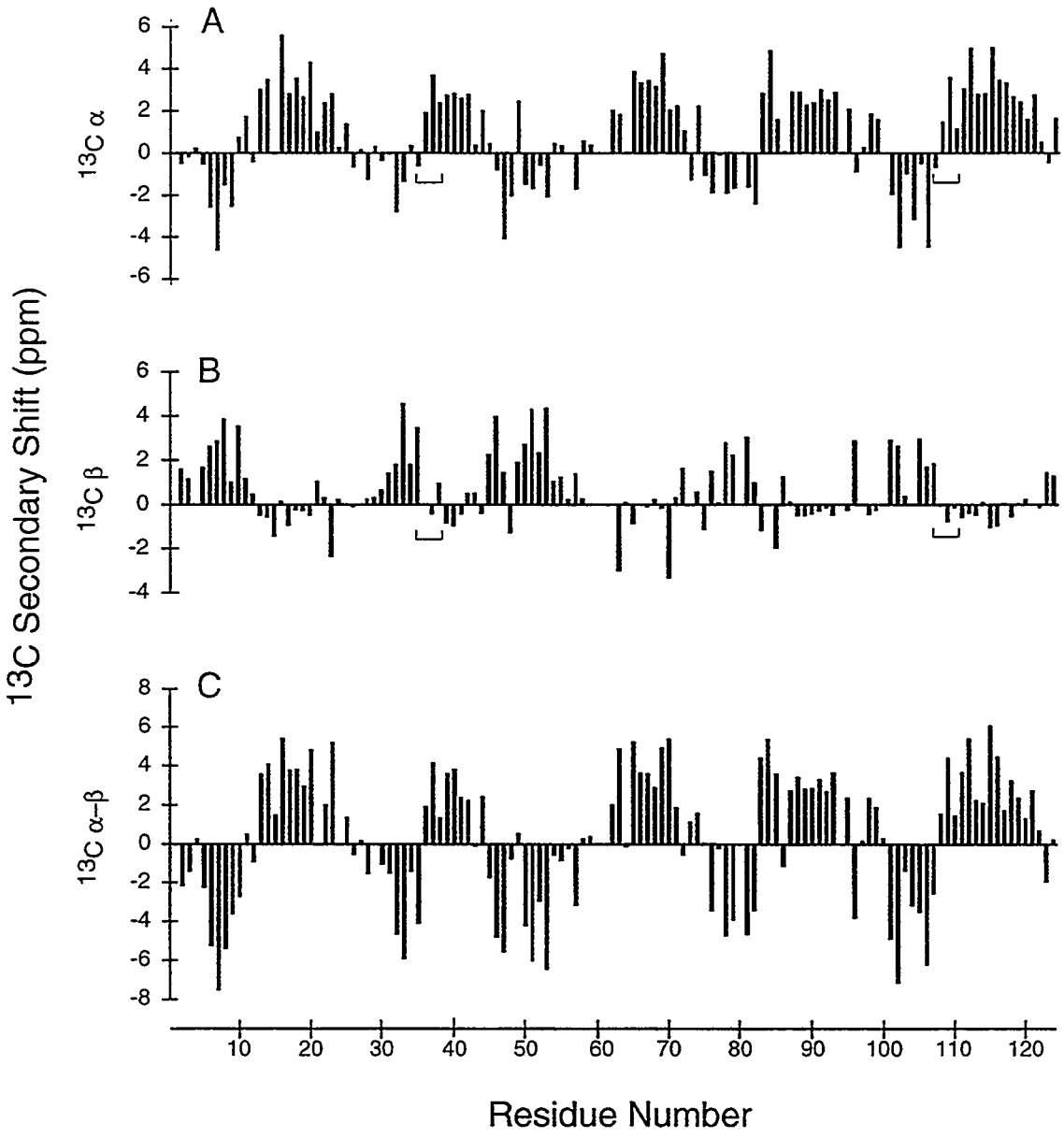


Figure 4.5: Observed ^{13}C secondary shifts of NTRC receiver domain, plotted as a function of residue number. Values reflect measured ^{13}C chemical shifts minus the amino acid-specific "random coil" ^{13}C chemical shift (Spera &Bax, 1991) in ppm, with positive and negative values indicating downfield and upfield secondary shifts, respectively. (A) ^{13}Cα secondary shifts. (B) ^{13}Cβ secondary shifts. (C) ^{13}Cα secondary shifts – ^{13}Cβ secondary shifts. Helical and extended regions are clearly evident in panel C as positive and negative regions, respectively. Patterns indicative of helix capping are indicated with brackets in panels A and B.

D49. In structure calculations without hydrogen bond restraints for these residues, the CO of I5 is reproducibly positioned so as to form hydrogen bonds with both D49 NH and V50 NH in a backbone conformation commonly characterized as a β-bulge (Richardson, 1981). At the same time, the sidechain CO of D49 is found to be the obvious acceptor for a hydrogen bond from I5 NH, appearing at an average distance of less than 3.0 Å. On the basis of these preliminary calculations, hydrogen bond restraints for these residues were included in the late rounds of refinement.

A β-bulge is also evident in strand 5. NOE patterns and amide exchange data indicate normal β-sheet hydrogen bonding between residues 81-102 and 82-103, but a lack of NOEs and amide protection for the preceding residues of strand 5 prevent the assignment of hydrogen bonds between 80-101 and 79-100. Additionally, the pattern of sequential NOEs for residues 98-100 is not consistent with extended β-sheet structure. NH-NH NOEs from 98 to 99 and 99 to 100, and a lack of strong sequential αH-NH NOEs reduce the likelihood of normal β-sheet formation for those residues. The NH of V78 is also protected from exchange, and the CO of A98 was determined to be its hydrogen bond acceptor. NOEs from the sidechains of F99 and Y101 position the residues of the bulge adjacent to strand 4, but without the backbone interactions typical of β-sheet structures.

Determination of the Three-Dimensional Structure of NTRC Receiver Domain: Table 4.2 summarizes the statistics of structure calculations of the NTRC receiver domain. Distance restraints for the structure calculations of NTRC receiver domains were generated from ^{15}N-edited and ^{13}C-edited spectra, as described in the Materials and Methods. A total of 915 experimental restraints were used, including 816 NOE-derived distance restraints, 19 dihedral angle restraints from the HMQC-J spectrum, and 80 restraints defining 40 hydrogen bonds. A total of 30 structures were calculated using the program X-PLOR 3.1. Of the 30 calculated structures, 20 with low final energies and minimal distance restraint violations were chosen for evaluation. Superposition of the residues contained in secondary structure, excluding helix 4, yields average root mean square deviations from the

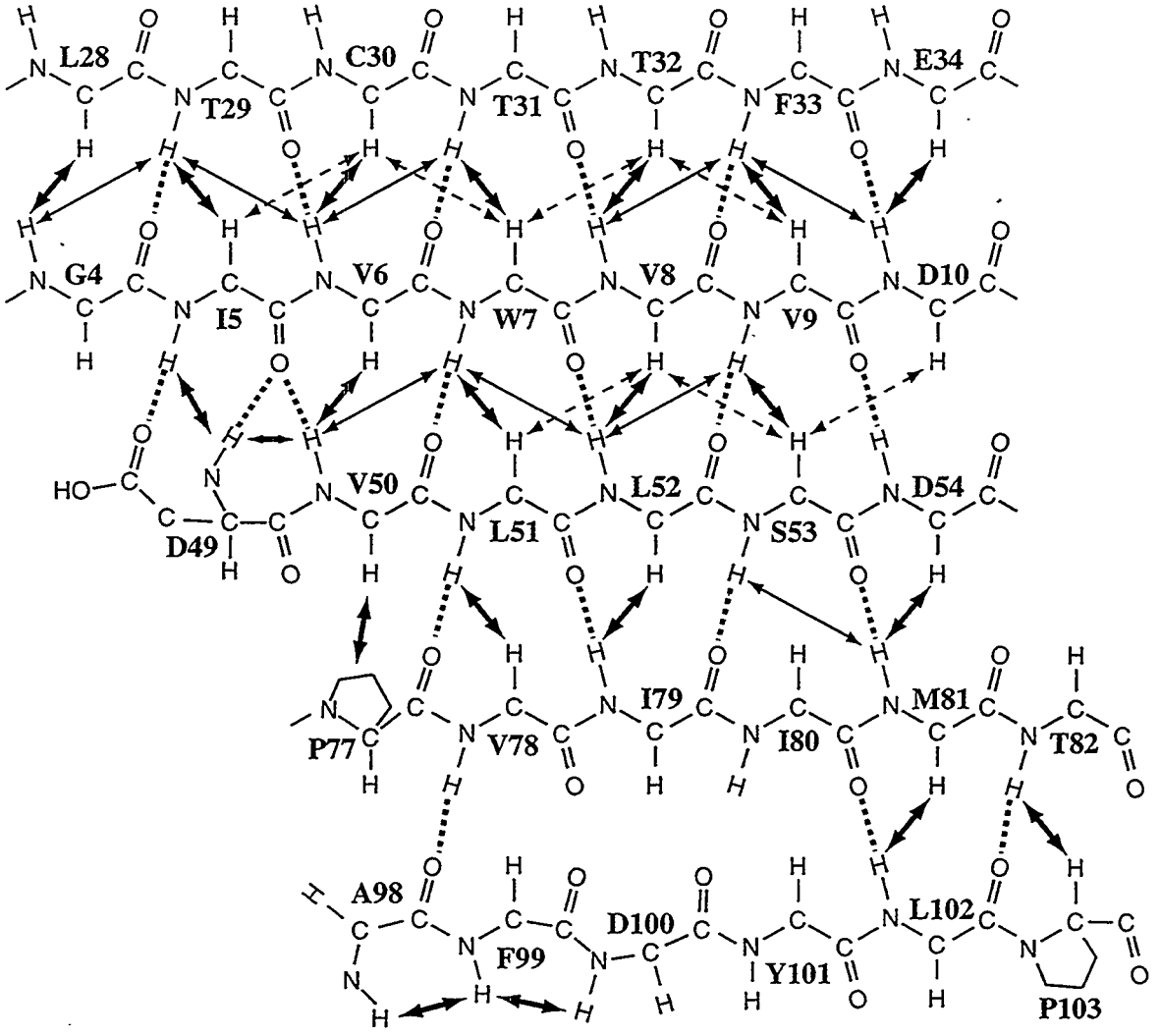


Figure 4.6: Cross-strand NOEs and hydrogen bonds observed for the β-sheet of NTRC receiver domain. Solid arrows correspond to NOEs observed in 3D ^{15}N NOESY-HMQC and dashed arrows correspond to NOEs observed in 2D ^{1}H-^{1}H NOESY. Dashed lines indicate hydrogen bonds included in structure calculations involving amide protons with reduced exchange rates.

average of 0.81 Å for the backbone atoms and 1.35 Å for non-hydrogen atoms. Figure 4.7 shows the family of 20 structures superimposed on the backbone atoms of the average structure. The ensemble reflects a single well-defined fold, with the loop from R56 to A64 being the only completely unstructured region.

The $(\beta/\alpha)_5$ fold has topological similarity to other α/β proteins (Richardson, 1981), with helices 1 and 5 nearly orthogonal to each other on one face of the sheet, and helices 2, 3 and 4 lying roughly parallel to each other on the other face. A number of hydrophobic interactions between the β-sheet and the helices are indicated by NOEs between sidechains of aliphatic and aromatic residues. One pocket of hydrophobic interactions involves residues V115 and I119 in helix 5, which have multiple NOE contacts to I79, L52 and V50. These interactions are important for defining the position of helix 5 next to the β-sheet. Unlike helices 1-4, helix 5 is not covalently constrained to the sheet at both ends and requires NOE restraints to define its position. Another group of sidechains including C30, L28, L23 and E20 is sufficiently buried to protect the sulfhydryl proton of C30 from rapid solvent exchange, allowing the normally unobserved resonance to be detected, even in experiments with presaturation of the solvent H_2O signal.

A ribbon diagram of the NTRC receiver domain is shown in Figure 4.8. Conserved residues of the active site form a cluster of sidechains at the C-terminal ends of the β-strands. The sidechain of D54 is the site of phosphorylation. The sidechains of residues D10, D54 and T82 are in close proximity due to their locations in the sheet. The ζNH_3^+ of K104 is oriented toward the sidechain of D54, but the complete degeneracy of the D54 $^1H\beta$ and $^{13}C\beta$ resonances with the K104 $^1H\varepsilon$ and $^{13}C\varepsilon$ resonances prevents the unambiguous assignment of NOEs which might position the K104 sidechain more precisely in the active site. The position of the sidechain of D11 is not well-defined in the family of structures due also to a lack of NOEs for that residue.

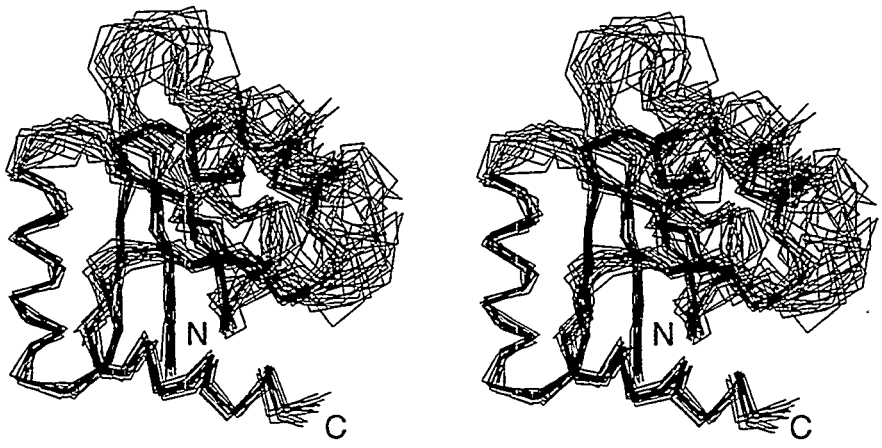


Figure 4.7: Stereoview of the family of 20 distance geometry-simulated annealing structures of NTRC receiver domain. Structures are superimposed on backbone atoms of the average structure, including residues 4-10, 14-44, 48-55, 65-73, 77-82, and 98-121.

Table 4.2 X-PLOR statistics for 20 NTRC receiver domain structures [a]

restraint totals by type	number
long range NOE	229
medium range NOE	180
sequential NOE	213
intraresidue NOE[b]	194
h-bond distance	82
dihedral	19
total restraints	917
X-PLOR energies (kcal/mol)	**<SA>**
E_{total}	322 ± 32
E_{bond}	13 ± 2
E_{angle}	189 ± 11
$E_{improper}$	25 ± 3
E_{vdw}[c]	42 ± 8
E_{noe}[d]	46 ± 11
E_{cdih}[e]	6.3 ± 1.7
RMSD from ideal geometry	**<SA>**
bonds (Å)	0.0026 ± 0.0002
angles (deg)	0.60 ± 0.02
impropers (deg)	0.41 ± 0.03
RMSD from experimental restraints	**<SA>**
distance restraints[f]	0.032 ± 0.004
dihedral restraints[f]	2.32 ± 0.30

Atomic RMSDs (Å)	N, Cα, C, O	all non-H
<SA> vs. $<\overline{SA}>_{2^\circ \text{ struct}}$	0.81 ± 0.06	1.35 ± 0.11
<SA> vs. $<\overline{SA}>_{\text{all residues}}$	1.50 ± 0.12	2.13 ± 0.09

[a]Notation is as follows:<SA> is the ensemble of 20 final X-PLOR structures. $<\overline{SA}>_{2^\circ \text{struct}}$ is the average coordinates for residues involved in secondary structure (4-10, 14-44, 48-55, 65-73, and 98-122) which were obtained from a least-squares superposition of those backbone (N, Cα, C, O) heavy atoms. $<\overline{SA}>_{\text{all residues}}$ is the average coordinates for residues 1-124 obtained from a least-squares superposition of those backbone heavy atoms. [b]Intraresidue restraints were included for NOEs between sidechain protons which were more than four bonds apart. [c]The X-PLOR F_{repel} function was used to simulate the van der Waals potential with atomic radii ranging from 0.9 times their CHARMM (Brooks et al., 1983) values at high tempetatures to 0.75 their CHARMM values at low temperatures (Brünger, 1992). [d]NOE-derived distance restaints were applied with a square-well potential with force constants of 50 kcal mol^{-1} $Å^{-2}$. [e]Dihedral angles were given force constants of 200 kcal mol^{-1} rad^{-2} which were applied at the beginning of the annealling/refinement stage. [f]A majority of NOE violations involved medium-range restraints in helix 4. A total of 2 NOE violations greater than 0.5 Å were found in the family of 20 accepted structures, and 1 dihedral restraint violation of greater than 6° was observed.

Discussion

The N-terminal receiver domain of the NTRC protein has been expressed at high levels and uniformly ^{15}N- and ^{13}C-labeled. The ^{1}H, ^{15}N, and ^{13}C resonance assignments have been completed using 3D ^{15}N- and ^{13}C-edited NMR techniques. Distance information was derived from 3D ^{15}N-edited NOESY-HMQC and 4D ^{13}C-edited HMQC-NOESY-HMQC spectra, while coupling constant and amide exchange information came from 2D ^{15}N-^{1}H experiments. The three-dimensional structure of the NTRC receiver domain was calculated using hybrid distance geometry/simulated annealing (DGSA) techniques. This structure provides a starting point from which to examine the effects of Mg^{2+} and phosphorylation on the NTRC receiver domain, and its subsequent interaction with the central domain of NTRC.

Comparison of NTRC receiver domain with CheY: High-resolution structures of CheY, a homologous receiver domain protein, have been determined by x-ray crystallography in the absence (Volz & Matsumura, 1991) and presence of Mg^{2+} (Stock et al., 1993; Bellsolell et al., 1994). Overall similarity between the NTRC receiver domain and CheY is high, as would be expected from the high degree of sequence conservation [29% identity for the proteins from enteric bacteria; (Volz, 1993)]. Superposition of only the residues of the β-sheet of the average NTRC receiver domain structure on each of the three high-resolution structures of CheY yields RMS deviations in Cα positions of 1.3 Å. All further comparisons of NTRC receiver domain to CheY were found to be identical for the three CheY structures.

There are two insertions in CheY relative to NTRC, but neither seems to have important structural consequences. Helix 3 of CheY has one extra turn at the C-terminus compared to helix 3 in the NTRC receiver domain, due to the presence of two additional residues in CheY in this region, and the termination of this helix in NTRC by P74. The other residue insertion in CheY relative to NTRC occurs between helix 1 and strand 2, just after the C-cap of helix 1, but has no significant effect upon the positioning of structural

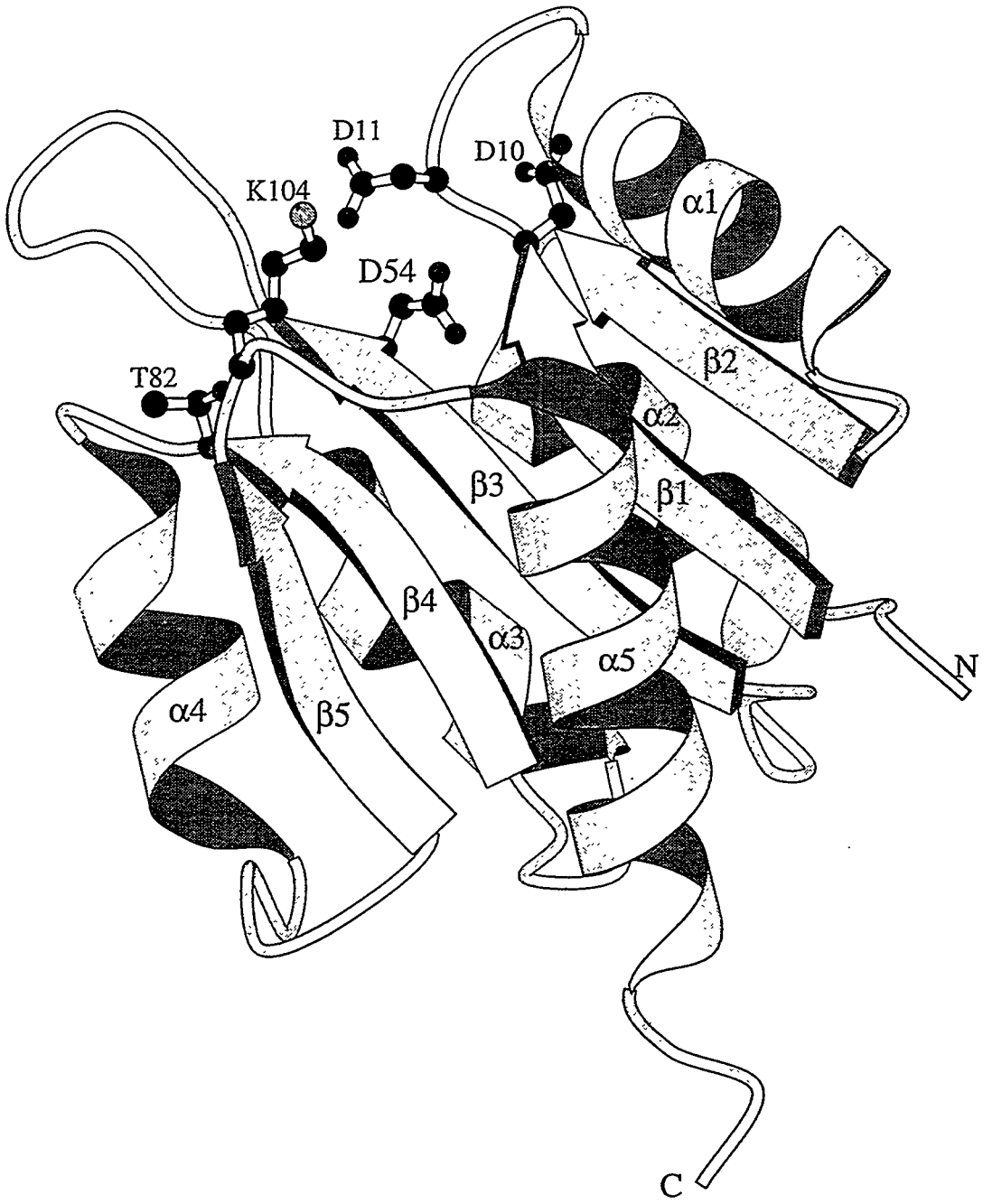


Figure 4.8: Ribbon diagram of the NTRC receiver domain. Secondary structure elements are individually labeled. Active site residues (D10, D11, D54, T82, K104) are labeled and shown in ball-and-stick representation.

elements. Orientations of helices 1, 2, 3, and 5 relative to the sheet are generally similar in the NTRC receiver domain and CheY. The Cα RMSD between the average NTRC structure and CheY is 2.7 Å for superposition of the sheet and helices 1, 2, 3, and 5.

Strikingly, inclusion of helix 4 in the superposition raises the RMSD value to 3.5 Å. When CheY and the NTRC receiver domain are superimposed on all secondary structural elements except helix 4, a difference in the orientation of the helix 4 axes of approximately 45° is observed. Figure 4.9A shows the superposition of the 20 DGSA structures of the NTRC receiver domain on the Cα trace of the crystal structure of CheY in the absence of Mg^{2+} (Brookhaven PDB file 3CHY). Average displacements from the corresponding CheY Cα coordinates of 8.1 to 9.5 Å are observed for the Cα atoms of residues L87 to A90 of NTRC. The RMS deviations of those four Cα atoms from the average NTRC receiver domain structure range from 1.3 to 1.8 Å, significantly smaller than the observed differences from CheY. Figure 4.9B illustrates in detail the difference in position of helix 4 in CheY and the family of NTRC structures. The range of coordinates spanned by residues 85-90 in the ensemble of NTRC structures clearly does not overlap the position of the same residues in CheY.

Active-site residues. The five conserved active site residues of the receiver domain superfamily are present in NTRC as well: D10, D11, D54, T82, and K104 (Moore et al., 1993). The resolution of this structure does not permit close comparison of these sidechains with the corresponding groups in CheY. However, the proximal positions of these residues in the structure of the NTRC receiver domain are consistent with their involvement in Mg^{2+} binding and phosphorylation.

Helix Capping . The Schellman C-terminal capping motif identified in helix 1 in the NTRC receiver domain is also present in the CheY structures. Examination of the sequence alignment for the receiver domain superfamily (Volz, 1993) reveals the conservation of the C' glycine, and the apolar residues at the C3 and C" positions, in accordance with the stereochemical rules for Schellman motifs (Aurora et al., 1994). The

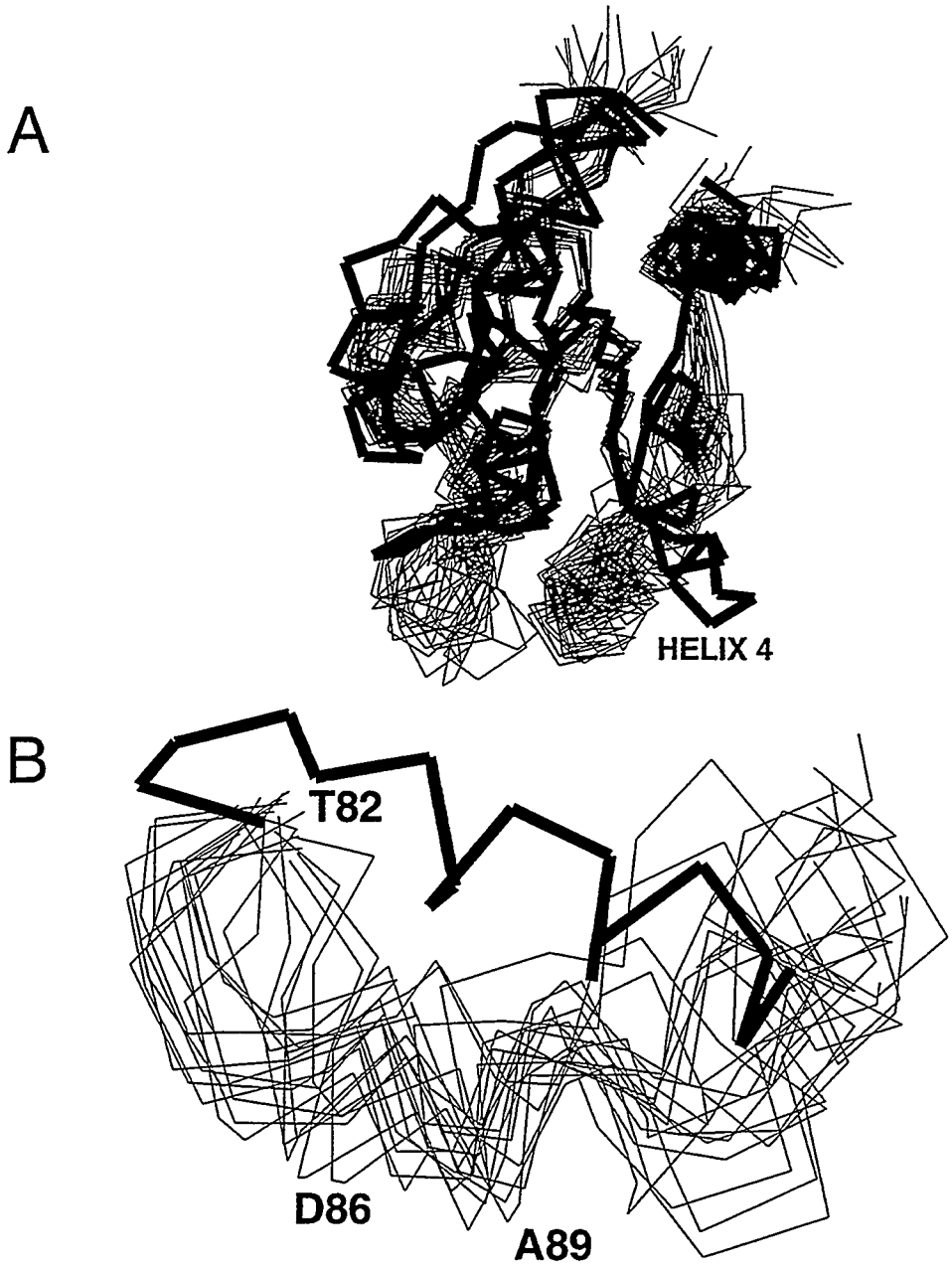


Figure 4.9: Superposition of 20 NTRC receiver domain structures on the high resolution crystal structure of CheY, demonstrating the difference in the position of helix 4 in the two proteins. The residues of all elements of secondary structure except for helix 4 were used for the alignment of structures. NTRC: 4-10, 15-45, 49-55, 66-72, 76-82, 97-122; CheY: 6-12, 17-30, 32-48, 52-58, 69-75, 81-87, 102-127. Cα traces of the same superposition of the full structures (A) and helix 4 in detail (B) are shown. The approximate locations of key residues from NTRC are indicated.

C3 position is a leucine in the consensus sequence, and the C" position is always an apolar residue, if the conserved glycine is present The solvent-exposed C1 position is a polar residue in nearly 90 % of the sequences. Mutation of key Schellman motif residues (C3, C', or C") can be very destabilizing, as seen in Staphylococcal nuclease (Shortle et al., 1990; Green et al., 1992). The conservation in the superfamily at these positions can be explained by the energetically favorable termination of helix 1 afforded by this C-capping motif.

The N-caps for helices 2 and 5 which are supported by the NMR data for the NTRC receiver domain are present in the crystal structures of CheY. The distances from the sidechain oxygen atoms of D38 and T112 to the amide nitrogens of D41 and T115 are 2.9Å and 3.2Å, respectively. No single residue type appears to be conserved at these capping positions in the CheY superfamily sequences. However, when all possible N-capping residues (S, T, D, N, E, Q, H, and C) are considered, a trend emerges. A potential N-cap for helix 2 is found in 86% of sequences, and in 65% of the sequences for helix 5. It is also interesting to consider the possibility of conserved "capping box" motifs (Harper & Rose, 1993). Conservation of capping box partners (N-cap and N3 position) is lower, but still significant: 59% for helix 2, and 46% for helix 5. Like the C-capping motifs, N-capping interactions provide an energetically favorable helix termination, forming one (N-cap) or two (capping box) additional hydrogen bonds which would otherwise be unsatisfied, and are a structural motivation for conservation at the N-cap and N3 positions.

β-bulges. The β-bulge and hydrogen bonding pattern at D49 in strand 3 of the NTRC receiver domain clarifies the basis for conservation at that position throughout the superfamily. The sidechain of D49 forms a cross-strand hydrogen bond to a backbone NH of strand 1, providing an additional stabilizing force at the N-terminal end of the β-sheet. This type of interaction may be present in other (α/β) proteins where aspartic acid is the most common N-terminal residue in β-strands (Colloc'h & Cohen, 1991). A similar

sidechain interaction is found in strand 5 in CheY, where a serine hydroxyl at the start of strand 5 accepts a hydrogen bond from a backbone amide in strand 4. The subsequent bulge is present in both NTRC and CheY, but the N-terminal residue of strand 5 in the NTRC receiver domain, A98, forms a normal backbone hydrogen bond to strand 4.

Function of helix 4 in the receiver domain of NTRC. As discussed above, helix 4 is the only structural element of the receiver domain of NTRC that is significantly repositioned with respect to CheY. Interestingly, two of the three "constitutive" amino acid substitutions so far identified in the receiver domain of NTRC (D86N and A89T) affect residues in helix 4 (Y. Flashner, J. Keener, and S. Kustu, unpublished results). The third substitution, D54E, affects the site of phosphorylation (Klose et al., 1993). NTRC constitutive proteins have some ability to activate transcription without being phosphorylated, both in vivo and in vitro. Hence, constitutive substitutions, which mimic phosphorylation, provide evidence for the functional importance of helix 4. It will be of interest to determine the relationship between structural changes in constitutive forms of the NTRC receiver domain and those that occur upon phosphorylation of the wild-type domain. The only constitutive substitutions known in CheY, D13K/R, appear to cause only local structural perturbations, whereas changes which occur upon phosphorylation of wild-type CheY are global (Bourret et al., 1993).

NMR Studies of the Mg^{2+} and phosphorylated forms of NTRC receiver domain: To more specifically address the issue of how the receiver domain of NTRC acts as a switch to turn on the ATPase of the central domain, further NMR structural studies are in progress. Because Mg^{2+} is required for phosphorylation of NTRC to occur, the structure of the NTRC receiver domain is to be determined in the presence of Mg^{2+}. The ultimate goal of this project is to structurally characterize the phosphorylated form of this receiver domain so that by comparison with the unphosphorylated structure we can observe the rearrangements which occur. This will provide a view of molecular signaling beyond levels of current understanding.

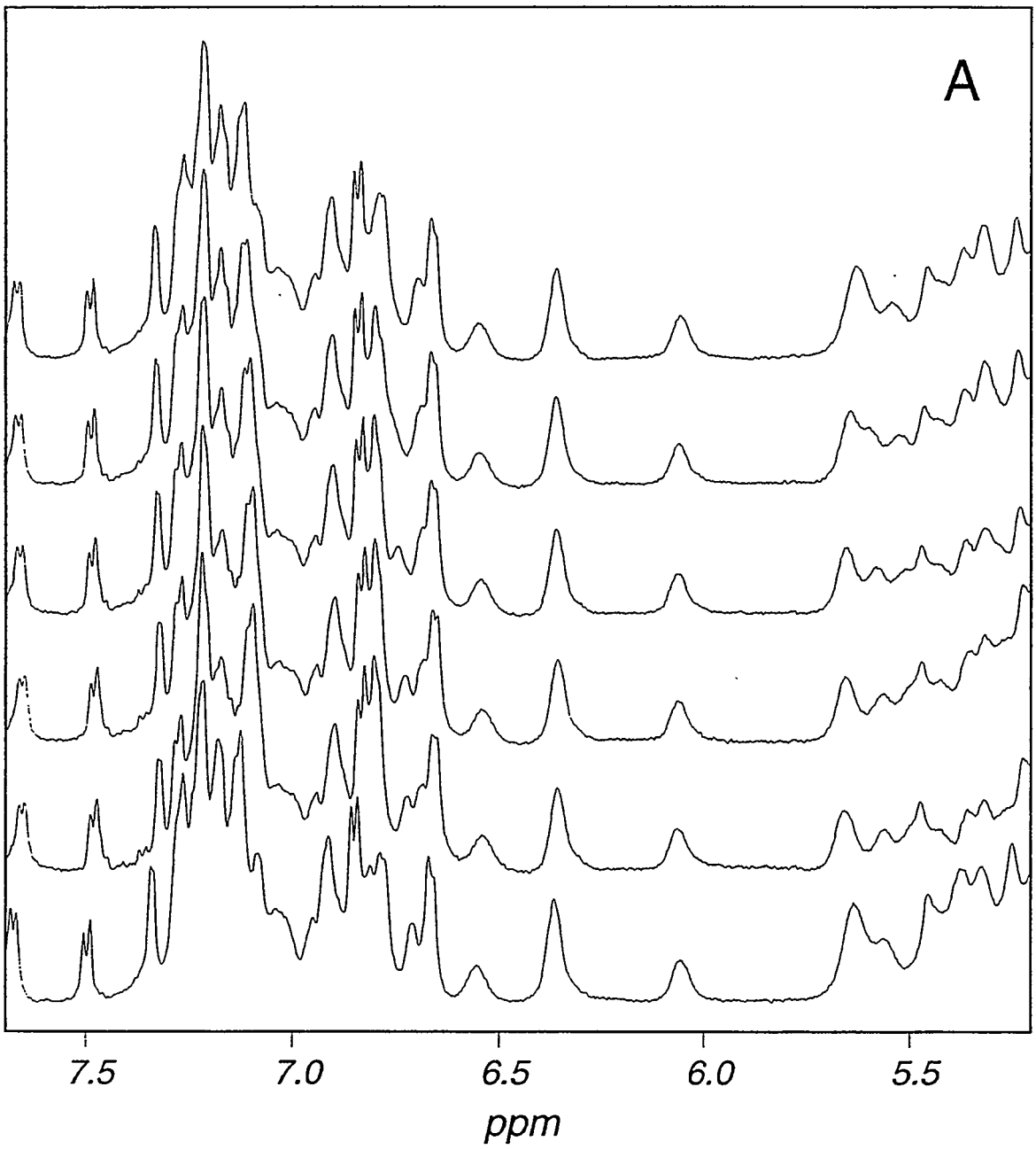
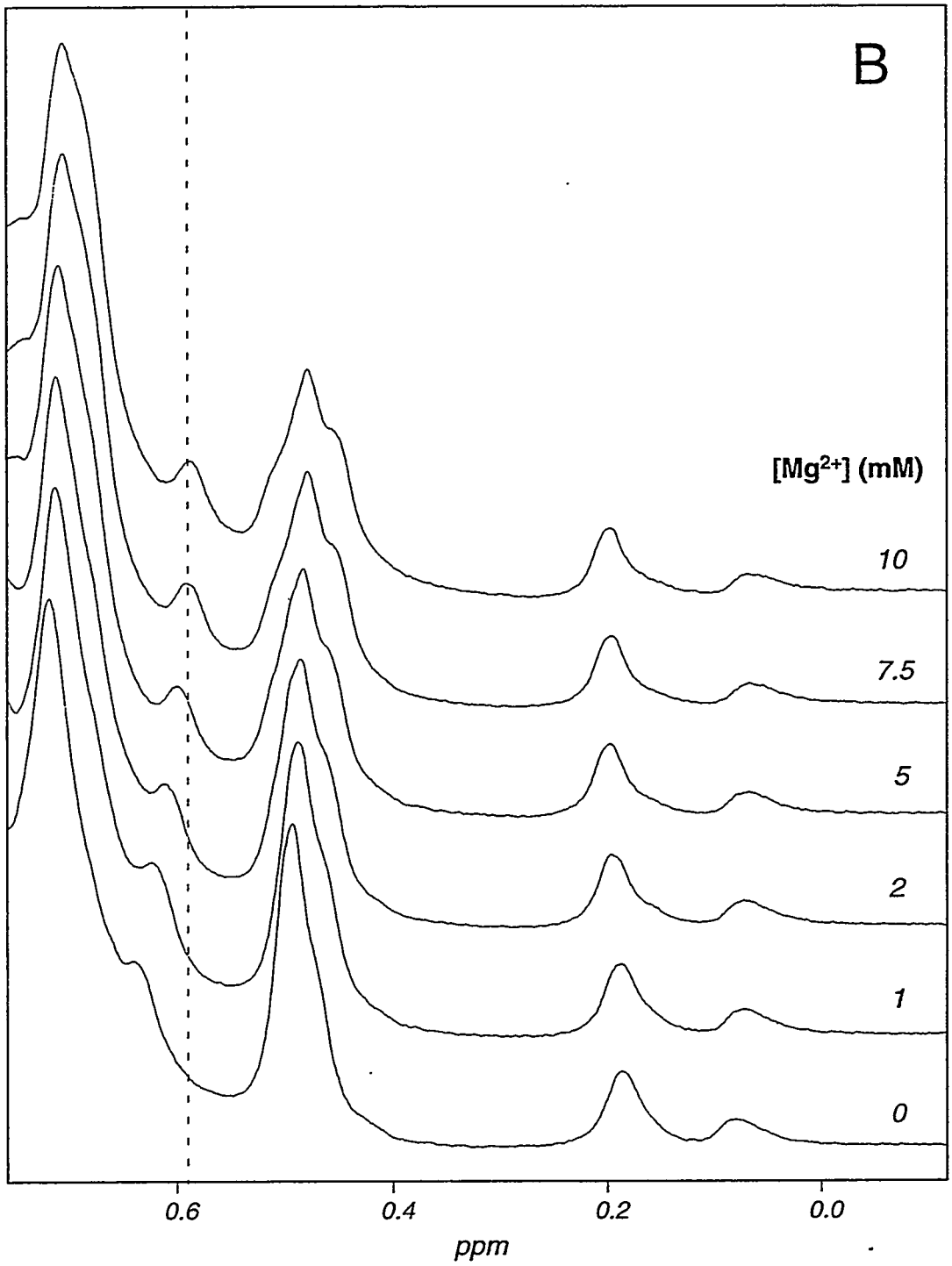


Figure 4.10: Titration of NTRC receiver domain with Mg^{2+}. Aromatic and αH resonances (panel A, this page) and upfield methyl resonances (panel B, next page) are shown, with increasing Mg^{2+} concentration corresponding to uppper traces as indicated in panel B.


B
[Mg2+] (mM)
10
7.5
5
2
1
0
0.6
0.4
0.2
0.0
ppm

Structural studies of CheY in the Mg^{2+}-bound and free forms reveal no significant changes in the positions of the backbone atoms of CheY upon binding of Mg^{2+}, so it is not anticipated that the structure of the NTRC receiver domain will be dramatically changed in the presence of Mg^{2+}. Figure 4.10 shows a titration of the NTRC receiver domain with Mg^{2+} monitored by ^{1}H NMR spectroscopy. A number of small shifts are observed in the aromatic, αH and methyl regions of the spectrum, but no large changes are apparent. Similar changes are observed in the ^{15}N-^{1}H HSQC spectrum as Mg^{2+} is added. At a Mg^{2+} concentration of between 5 mM and 10 mM precipitation of protein is observed in the NMR tube. This precipitation is observed at a pH of both 6.4 and 7.0 in both phosphate and TRIS buffers.

Phosphorylation of the NTRC receiver domain is accomplished by the addition of either acetyl phosphate or carbamyl phosphate in the presence of Mg^{2+}. Figure 4.11 shows ^{1}H NMR spectra of the NTRC receiver domain demonstrating the effect of carbamyl phosphate upon resonances in the αH region of the spectrum. The most obvious change is the appearance of a new peak at 5.9 ppm. The changes effected by carbamyl phosphate are Mg^{2+}-dependent as new peaks do not appear without Mg^{2+}, and the addition of EDTA to a phosphorylated sample results in the fast recovery of the original spectrum due to the autophosphatase activity of NTRC. The questions to be addressed before attempting to perform 3D and 4D NMR experiments on phosphorylated NTRC receiver domain are twofold. First, is a steady-state level of phosphorylation of nearly 100% achieved with carbamyl phosphate (or the as-yet-untried phosphoramidate), and if so, how long can that steady-state be maintained? Indications are that by adding 50-100 mM carbamyl phosphate a high level of phosphorylation can be maintained for a timescale of hours, but a maximum duration has not been determined. Inspection of 2D HSQC measurments indicates that a very high percentage of protein is being phosphorylated, with unphosphorylated protein being reduced to undetectable levels in some instances. Secondly, is the protein being phosphorylated specifically at D54 and nowhere else? To test this, mutants (D54E and

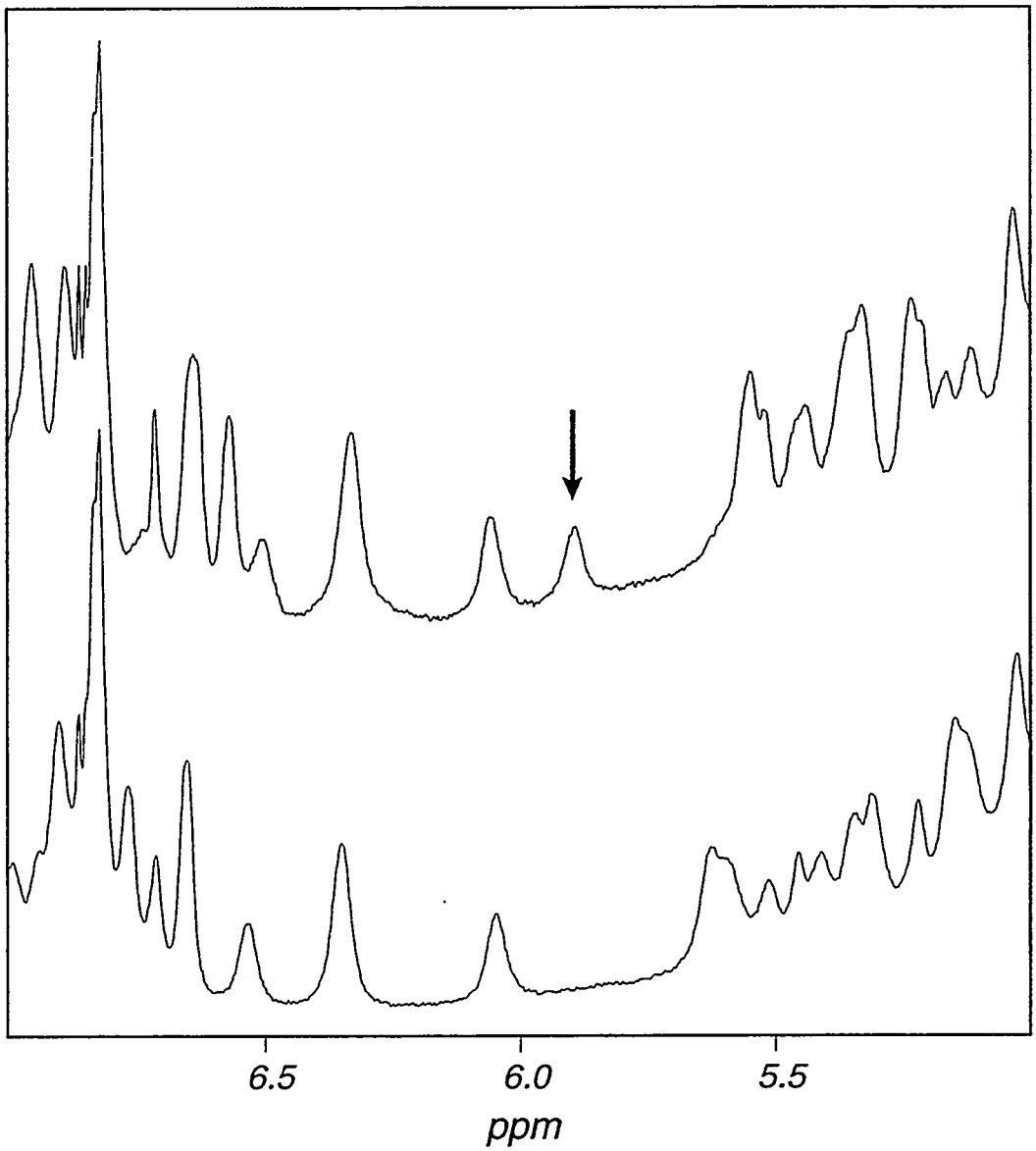


Figure 4.11: Phosphorylation of the NTRC receiver domain with carbamyl phosphate. NMR spectra were collected at 25° C in D_2O in the presence of 10 mM Mg^{2+}. The upper spectrum was collected 1 hour after addition of 50 mM carbamyl phosphate to the sample shown in the lower trace. The appearance of a new peak at 5.8 ppm (indicated with arrow) signals the production of the phosphorylated form of the NTRC receiver domain.

D54A) which are not phosphorylatable by NTRB will be assayed by NMR to see if any spectral changes occur upon addition of Mg^{2+} and carbamyl phosphate. It is not expected that non-specific phosphorylation should occur at significant levels, but it is relatively simple to check, and might complicate assignment of NMR spectra if other forms are detected.

Bibliography

Anglister, J., Frey, T. & McConnell, H. M. (1984) *Biochemistry 23*, 1138-1142.

Anglister, J., Grzesiek, S., Ren, H., Klee, C. B. & Bax, A. (1993) *J. Biol. NMR 3*, 121-126.

Anil-Kumar, Ernst, R. R. & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1-6.

Aurora, R., Srinivasan, R. & Rose, G. D. (1994) *Science 264*, 1126-1130.

Austin, S. & Dixon, R. (1992) *EMBO J. 11*, 2219-2228.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1987) *Current Protocols in Molecular Biology*, (K. Janssen), John Wiley & Sons, New York.

Bagg, A. & Neilands, J. B. (1987) *Biochemistry 26*, 5471-5477.

Barany, G. & Merrifield, B. (1980) in *The Peptides* (E. Gross and J. Meienhofer, ed.) 1-284, Academic Press, New York.

Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W. & Bax, A. (1992) *Biochemistry 31*, 5269-5278.

Bax, A., Clore, G. M. & Gronenborn, A. M. (1990) *J. Magn. Reson. 88*, 425-431.

Bax, A. & Davis, G. (1985) *J. Magn. Reson. 65*, 355-360.

Bellsolell, L., Prieto, J., Serrano, L. & Coll, M. (1994) *J. Mol. Biol. 238*, 489-495.

Bodenhausen, G. & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185-189.

Bourret, R. B., Drake, S. K., Chervitz, S. A., Simon, M. I. & Falke, J. J. (1993) *J. Biol. Chem. 268*, 13089-96.

Braunschweiler, L. & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521-528.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comput. Chem. 4*, 187-217.

Brünger, A. T. (1992) *X-PLOR Manual Version 3.1*, Yale University Press, New Haven, CT.

Burum, D. P. & Ernst, R. R. (1980) *J. Magn. Reson. 39*, 163.

Chang, C., Kwok, S. F., Bleecker, A. B. & Meyerowitz, E. M. (1993) *Science 262*, 539-544.

Clore, G. M. & Gronenborn, A. M. (1991) *Science 252*, 1390-1399.

Clore, G. M., Kay, L. E., Bax, A. & Gronenborn, A. M. (1991) *Biochemistry 30*, 12-18.

Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M. & Gronenborn, A. M. (1986) *EMBO J. 5*, 2729-2735.

Colloc'h, N. & Cohen, F. E. (1991) *J. Mol. Biol. 221*, 603-613.

Coy, M. (1993) *PhD Thesis*, U. C. Berkeley,

Coy, M. & Neilands, J. B. (1991) *Biochemistry 30*, 8201-8210.

Davies, D. R., Sheriff, S. & Padlan, E. A. (1988) *J. Biol. Chem. 263*, 10541-10544.

Davis, D. G. & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820-2821.

de Lorenzo, V., Giovannini, F., Herrero, M. & Neilands, J. B. (1988a) *J. Mol. Biol. 203*, 875-884.

de Lorenzo, V., Herrero, M., Giovannini, F. & Neilands, J. B. (1988b) *Eur. J. Biochem. 173*, 537-546.

de Lorenzo, V., Wee, S., Herrero, M. & Neilands, J. B. (1987) *J. Bacteriol. 169*, 2624-2630.

Defendini, M.-L., Pierres, M., Regnier-Vigouroux, A., Rochat, H. & Granier, C. (1990) *Molecular Immunology 27*, 551-558.

Doolittle, R. F. (1986) *Of URFS and ORFS: A primer on how to analyze derived amino acid sequences*, University Science Books, Mill Valley, CA.

Dorman, D. E. & Bovey, F. A. (1973) *J. Org. Chem. 38*, 2379-2383.

Driscoll, P. C., Clore, G. M., Marion, D., Wingfield, P. T. & Gronenborn, A. M. (1990) *Biochemistry 29*, 3542-3556.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D. & Wemmer, D. (1979) *Symp. Faraday Soc. 13*, 49-55.

Drummond, M. H., Conteras, A. & Mitchenall, L. A. (1990) *Molecular Microbiology 4*, 29-37.

Dyson, H. J., Lerner, R. A. & Wright, P. E. (1988) *Ann. Rev. Biophys. Biophys. Chem. 17*, 305-324.

Ernst, R. R., Bodenhausen, G. & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, (R. B. Bernstein, R. Breslow, M. L. H. Green, J. Halpern and J. S. Rowlinson), Clarendon Press, Oxford.

Feng, J., Atkinson, M., McCleary, W., Stock, J., Wanner, B. & Ninfa, A. (1992) *J. Bacteriol. 174*, 6061-6070.

Gill, S. C. & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319-326.

Green, S. M., Meeker, A. K. & Shortle, D. (1992) *Biochemistry 31*, 5717-5728.

Griesinger, C., Otting, G., Wüthrich, K. & Ernst, R. R. (1988) *J. Magn. Reson. 110*, 7870-7820.

Gronenborn, A. M. & Clore, G. M. (1994) *J. Biomol. NMR 4*, 455-458.

Grzesiek, S. & Bax, A. (1992a) *J. Am. Chem. Soc. 114*, 6291-6293.

Grzesiek, S. & Bax, A. (1992b) *J. Magn. Reson. 99*, 201-207.

Halliwell, B. (1988), *Oxygen Radicals and Tissue Injury,* Federation of the American Society for Experimental Biology, Bethesda, MD.

Harper, E. T. & Rose, G. D. (1993) *Biochemistry 32*, 7605-7609.

Harris, R. K. (1986) *Nuclear Magnetic Resonance Spectroscopy*, Longman Scientific & Technical, New York.

Havel, T., Kuntz, I. D. & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665-720.

Havel, T. & Wüthrich, K. (1984) *Bull. Math. Biol. 46*, 673-698.

Haynes, M. R., Stura, E. A., Hilvert, D. & Wilson, I. A. (1994) *Science 264*, 646.

Hess, J. F., Oosawa, K., Kaplan, N. & Simon, M. I. (1988) *Cell 53*, 79-87.

Hicks, R. P., Young, J. K. & Moskau, D. (1994) *Concepts in Magnetic Resonance 6*, 115-130.

Kay, L. E. & Bax, A. (1990) *J. Magn. Reson. 86*, 110-126.

Kay, L. E., Ikura, M., Tschudin, R. & Bax, A. (1990) *J. Magn. Reson. 89*, 496-514.

Kay, L. E., Marion, D. & Bax, A. (1989) *J. Magn. Reson. 84*, 72-84.

Keener, J. & Kustu, S. (1988) *Proc. Natl. Acad. Sci. USA 85*, 4976-4980.

King, D. S., Fields, C. G. & Fields, G. B. (1990) *Int. J. Peptide Protein Res. 36*, 255-266.

Klausner, R. D., Rouault, T. A. & Harford, J. B. (1993) *Cell 72*, 19.

Klose, K., North, A., Stedman, K. & Kustu, S. (1994) *J. Mol. Biol. 240*, in press.

Klose, K. E., Weiss, D. W. & Kustu, S. (1993) *J. Mol. Biol. 232*, 67-78.

Kohler, G. & Milstein, C. (1975) *Nature 256*, 495-497.

Kraulis, P. J. (1991) *J. App. Cryst. 24*, 946-950.

Kumar, A., Ernst, R. R. & Wüthrich, K. (1980) *Biochem. Biophys. Res. Com. 95*, 1-6.

Kuszewski, J., Nilges, M. & Brünger, A. T. (1992) *J. Biomol. NMR 2*, 33-56.

LeMaster, D. M. (1990) *Quarterly Reviews of Biophysics 23*, 133-174.

Lerner, R. A., Benkovic, S. J. & Schultz, P. G. (1991) *Science 252*, 659.

Live, D. H., Davis, D. G., Agosta, W. C. & Cowburn, D. (1984) *J. Am. Chem. Soc. 106*, 1934-1941.

Lukat, G. S., Lee, B. H., Mottonen, J. M., Stock, A. M. & Stock, J. B. (1991) *J. Biol. Chem. 266*, 8348-8354.

Lukat, G. S., McCleary, W. R., Stock, A. M. & Stock, J. B. (1992) *Proc. Natl. Acad. Sci. USA 89*, 718-722.

Lyu, P. C., Wemmer, D. E., Zhou, H. X., Pinker, R. J. & Kallenbach, N. R. (1993) *Biochemistry 32*, 421-425.

Maeda, T., Wurgler-Murphy, S. M. & Saito, H. (1994) *Nature 369*, 242-245.

Manoleras, N. & Norton, R. S. (1992) *J. Biomol. NMR 2*, 485-494.

Marion, D., Ikura, M., Tschudin, R. & Bax, A. (1989a) *J. Magn. Reson. 85*, 393-399.

Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A. & Bax, A. (1989b) *J. Am. Chem. Soc. 111*, 1515-1517.

McIntosh, L. P. & Dahlquist, F. W. (1990) *Quarterly Reviews of Biophysics 23*, 1-38.

Moore, J. B., Shiau, S. P. & Reitzer, L. J. (1993) *J. Bacteriol. 175*, 2692-2701.

Morris, G. A. & Freeman, R. (1979) *J. Am. Chem. Soc. 101*, 760.

Neilands, J. B. (1981) *Ann. Rev. Biochem. 50*, 715-731.

Neilands, J. B. (1990a) in *Pseudomonas: Biotransformation, Pathogenesis, and Evolving Biotechnology* ed.) 382-395, American Society for Microbiology, Washington, D. C.

Neilands, J. B. (1990b) *Adv. Inorg. Biochem 8*, 63-90.

Nilges, M., Clore, G. M. & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317-324.

Ninfa, A. J. & Magasanik, B. (1986) *Proc. Natl. Acad. Sci. USA 83*, 5909-5913.

Noggle, J. H. & Schirmer, R. E. (1971) *The nuclear Overhauser effect - chemical applications*, Academic Press, New York.

Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N. & Campbell, I. D. (1990) *J. Magn. Reson. 87*, 488-501.

Ota, I. M. & Varshavsky, A. (1993) *Science 262*, 566-569.

Parkinson, J. S. & Kofoid, E. C. (1992) *Ann. Rev. Genet. 26*, 71-112.

Pease, J. H. B., Storrs, R. W. & Wemmer, D. E. (1990) *Proc. Natl. Acad. Sci. USA 87*, 5643-5647.

Pease, J. H. B. & Wemmer, D. E. (1988) *Biochemistry 27*, 8491-8498.

Porter, S. C. (1993) *PhD Thesis* U. C. Berkeley.

Porter, S. C., North, A. K. & Kustu, S. (in press) in *Signal Transduction in Bacteria* (T. Silhovy and T. Hoch, ed.)

Presta, L. G. & Rose, G. D. (1988) *Science 240*, 1632-1641.

Ravid, S., Matsumura, P. & Eisenbach, M. (1986) *Proc. Natl. Acad. Sci. USA 83*, 7157-7161.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167-337.

Saito, T. & Williams, R. J. P. (1991) *Eur. J. Biochem. 197*, 43-47.

Saito, T., Wormald, M. R. & Williams, R. J. P. (1991) *Eur. J. Biochem. 197*, 29-38.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Sanders, D. A., Gillece-Castro, B. L., Burlingame, A. L. & Koshland, D. E., Jr. (1992) *J. Bacteriol. 174*, 5112-5122.

Scanlon, M. J. & Norton, R. S. (1994) *Protein Science 3*, 1121-1124.

Schellman, C. (1980) in *Protein Folding* (R. Jaenicke, ed.) 53-61, Elsevier, North-Holland, New York.

Shaka, A. J., Lee, C. J. & Pines, A. (1988) *J. Magn. Reson. 77*, 274-293.

Shortle, D., Stites, W. E. & Meeker, A. K. (1990) *Biochemistry 29*, 8033-8041.

Sklenar, V. & Bax, A. (1987) *J. Magn. Reson. 74*, 469-479.

Spera, S. & Bax, A. (1991) *J. Am. Chem. Soc. 113*, 5490-5492.

Stock, A. M., Martinez-Hackert, E., Rasmussen, B. F., West, A. H., Stock, J. B., Ringe, D. & Petsko, G. A. (1993) *Biochemistry 32*, 13375-13380.

Stock, A. M., Mottonen, J. M., Stock, J. B. & Schutt, C. E. (1989) *Nature 337*, 745-749.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60-89.

Sutcliffe, J. G., Shinninck, T. M., Green, N., Liu, F.-T., Niman, H. L. & Lerner, R. A. (1980) *Nature 287*, 801-805.

Volz, K. (1993) *Biochemistry 32*, 11741-11753.

Volz, K. & Matsumura, P. (1991) *J. Biol. Chem. 296*, 15511-15519.

Vuister, G. W. & Bax, A. (1992) *J. Magn. Reson. 98*, 428-435.

Walter, G., Scheidtmann, K.-H., Carbone, A., Laudano, A. P. & Doolittle, R. F. (1980) *Proc. Natl. Acad. Sci. USA 77*, 5197-5200.

Wee, S., Neilands, J. B., Bittner, M. L., Hemming, B. C., Haymore, B. L. & Seetharam, R. (1988) *Biol. Metals 1*, 62-68.

Weiss, D. S., Batut, J., Klose, K. E., Keener, J. & Kustu, S. (1991) *Cell 67*, 155-167.

Weiss, D. S., Klose, K. E., Hoover, T. R., North, A. K., Porter, S. C., Wedel, A. B. & Kustu, S. (1992) in *Transcriptional Regulation* (S. L. McKnight and K. R. Yamamoto, ed.) 667-694, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Weiss, V., Claverie-Martin, F. & Magasanik, B. (1992) *Proc. Natl. Acad. Sci. USA 89*, 5088-5092.

Weiss, V. & Magasanik, B. (1988) *Proc. Natl. Acad. Sci. USA 85*, 8919-8923.

Williamson, M. P., Havel, T. F. & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295-315.

Wüthrich, K. (1976) *NMR in biological research: Peptides and proteins*, North Holland, Amsterdam.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.

Wüthrich, K., Billeter, M. & Braun, W. (1983) *J. Mol. Biol. 169*, 949-961.

Zhou, G. W., Guo, J., Huang, W., Fletterick, R. J. & Scanlan, T. S. (1994a) *Science 265*, 1059-1064.

Zhou, H. X., Lyu, P. C., Wemmer, D. E. & Kallenbach, N. R. (1994b) *Proteins: Struct., Funct. Genet. 18*, 1-7.

Zhu, G. & Bax, A. (1990) *J. Magn. Reson. 90*, 405.

Appendix A
Supplementary material to chapter 1

1. Excel macro for 4D NOESY assignment

The extract3.mac macro sheet is for use with a set of three other Excel worksheets. The extract3 macro reads the chemical shifts for each 4D NOESY crosspeak one at a time from the "see4dppm.wsh" sheet, which also contains the criteria for matching with the database of chemical shifts assignments. These criteria are used for an 'extract' operation, which is a powerful Excel command that automatically searches a predefined 'database' and reports all entries which satisfy the criteria. The database is contained in the second worksheet, and must be defined internally in the extract3.mac macro sheet (i.e., it is never explicitly referenced in the extract3.mac macro shown below, only implicitly in the actual extract(true) command). The results of the extract command are deposited in a third sheet, in this case called extract2.wsh. At the completion of this macro, the extract2.wsh contains all of the possible matches for both the starting and destination proton-carbon pairs for each 4D NOESY crosspeak, as well as the offset (error) in each of the matching condition. This sheet can then be examined manually to obtain distance restraints for structure calculations, or used as input for semi-automated restraint generation programs (for comparison with low-resolution initial structures, etc.).

extract3.mac

```
=ECHO(FALSE)                                          turn off screen updating
=FOR("inc",1,1617,1)                                  2nd and 3rd arguments
                                                      define loop
=ACTIVATE("see4dppm.wsh")                             this appears to be necessary
=DEFINE.NAME("incells",OFFSET(!$B$10,inc,0,1,3))
=DEFINE.NAME("incell",OFFSET(!$B$10,inc,0,1,1))
=ACTIVATE("extract2.wsh")
=SELECT.SPECIAL(11)                                   select last cell
=DEFINE.NAME("lastcell")
=DEFINE.NAME("top2",OFFSET(ACTIVE.CELL(),1,-3))       1
=COPY('see4dppm.wsh'!incells,OFFSET(ACTIVE.CELL(),1, copy  #,sign,intensity
-25))
=COPY("exhead",OFFSET(ACTIVE.CELL(),1,-21))
=COPY("exhead",OFFSET(ACTIVE.CELL(),1,-11))
                                                      now ready to set criteria
=ACTIVATE("see4dppm.wsh")
=SELECT("incell")
=SET.NAME("d1prot",OFFSET(ACTIVE.CELL(),0,3))         get proton (D1) value
=SET.NAME("d2prot",OFFSET(ACTIVE.CELL(),0,4))
=SET.NAME("d3carb",OFFSET(ACTIVE.CELL(),0,5))
=SET.NAME("d4carb",OFFSET(ACTIVE.CELL(),0,6))         get carbon (D4) value
=SELECT(OFFSET(ACTIVE.CELL(),0,9,1,4))
=COPY(SELECTION(),!$K$7:$N$7)                         this updates the criteria
```

```
=SELECT(OFFSET(ACTIVE.CELL(),0,4,1,4))
=COPY(SELECTION(),!$K$2:$N$2)                          copies other half of noe
=COPY(!$K$7:$N$7,!$K$8:$N$8)
=COPY(!$K$7:$N$7,!$K$9:$N$9)
=COPY(!$K$2:$N$2,!$K$3:$N$3)
=COPY(!$K$2:$N$2,!$K$4:$N$4)                           copies down
=SELECT(!$K$8:$N$8)
=FORMULA.REPLACE("fold2","fold2+18.614",,,FALSE)       account for folding
=SELECT(!$K$9:$N$9)
=FORMULA.REPLACE("fold2","fold2+37.228",,,FALSE)       account for folding
=SELECT(!$K$3:$N$3)
=FORMULA.REPLACE("fold2","fold2+18.614",,,FALSE)       other half of noe
=SELECT(!$K$4:$N$4)
=FORMULA.REPLACE("fold2","fold2+37.228",,,FALSE)

=ACTIVATE("extract2.wsh")
=SELECT("lastcell")
=SELECT(OFFSET(ACTIVE.CELL(),1,-21,1,10))
=SET.EXTRACT()                                         set extract range
=EXTRACT(TRUE)
=SET.NAME("prot",d1prot)
=SET.NAME("carb",d4carb)
=SET.NAME("round",1)
=GOTO(difference)

back1
=SELECT(OFFSET(!top2,0,-8,1,10))
=SET.EXTRACT()                                         reset extract range
=COPY('see4dppm.wsh'!$K$2:$N$4,'see4dppm.wsh'!$K$      copy new criteria into range
7:$N$9)
=EXTRACT(TRUE)
=SET.NAME("prot",d2prot)
=SET.NAME("carb",d3carb)
=SET.NAME("round",2)
=GOTO(difference)

back 2
=SELECT.LAST.CELL()                                    fill remaining empty spaces
=SELECT(OFFSET(!lastcell,1,-
15):OFFSET(ACTIVE.CELL(),0,-15))
=FOR.CELL("fill")
=IF(ISBLANK(fill))
=FORMULA("-",fill)
=ELSE()
=SET.NAME("dummy",1)
=END.IF()
=NEXT()

=SELECT.LAST.CELL()                                    fill remaining empty spaces
=SELECT(OFFSET(!lastcell,1,-
5):OFFSET(ACTIVE.CELL(),0,-5))
=FOR.CELL("fill")
```

```
=IF(ISBLANK(fill))
=FORMULA("-",fill)
=ELSE()
=SET.NAME("dummy",1)
=END.IF()
=NEXT()

=NEXT()
=ECHO(TRUE)                                          turn on screen updating
=RETURN()                                            end of macro

difference                                           subroutine to calc error
                                                     offset
=SELECT(OFFSET(ACTIVE.CELL(),0,6))                   go to 'proton'
=DEFINE.NAME("top")

=SELECT(OFFSET(ACTIVE.CELL(),-1,0))                  go one above 'proton'
=SELECT.END(4)                                       select last cell
=SET.NAME("temp",ROW(ACTIVE.CELL())-                 set length of loop
ROW(!lastcell))
=SET.NAME("length",temp-1)

=IF(length=0,GOTO(redirect))                         no calc if none extracted

=FOR("notch",1,length,1)
=SELECT(OFFSET(!top,notch,0))
=FORMULA(GET.CELL(5,ACTIVE.CELL())-                  enters difference next to
prot,OFFSET(ACTIVE.CELL(),0,1))                      proton column
=FORMULA(GET.CELL(5,OFFSET(ACTIVE.CELL(),0,2))-      same for carbon
carb,OFFSET(ACTIVE.CELL(),0,3))
=SELECT(OFFSET(ACTIVE.CELL(),0,3))                   go to carbon error cell

=IF(ABS(GET.CELL(5,ACTIVE.CELL()))<10)               if no folding ...
=SET.NAME("dummy",1)
=ELSE.IF(GET.CELL(5,ACTIVE.CELL())>10)
=FORMULA(GET.CELL(5,ACTIVE.CELL())-                  subtract 20.45
18.614,ACTIVE.CELL())
=ELSE()
=FORMULA(GET.CELL(5,ACTIVE.CELL())+18.614,ACTIVE     add 20.45
.CELL())
=END.IF()

=NEXT()

redirect
=IF(round=1)                                         loop back
=GOTO(back1)
=ELSE()                                              then round must = 2
=GOTO(back2)
=END.IF()
```

2. AMX-600 pulse programs

```
;zgpr
;1D sequence with presaturation


1 ze
2 d12 hl2
  p18 ph29
  d13
  d12 hl1
  p1 ph1
  go=2 ph31
  wr #0
exit


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


;DP=hl1: ecoupler high power level
;hl2: ecoupler power level for presaturation
;p1 :  90 degree transmitter high power
pulse
;p18: presaturation during relaxation delay
;d12: delay for power switching    [20 usec]
;d13: short delay (e.g. to compensate
                          delay line)    [3 usec]
```

```
;mlevcprtp.ok
;homonuclear Hartman-Hahn transfer using
;compensated MLEV17 sequence for mix
;one power level for excitation and SL
;phase sensitive using TPPI
;d17 delays in mixing period suppress
;effects of cross-relaxation
;C. Griesinger, et al. JACS, 110:7870
;A. Bax * D. G. Davis, J. Magn. Reson.
;65, 355-360 (1985)
;with MLEV17y instead of MLEV17x

1 ze
2 d11
3 d12 hl2
  d6
  p18 ph29
  d13
  d12 hl3
  p1 ph1
  d0
  (p17 ph26)
4 (p6 ph22 d17 p7 ph23 d17 p6 ph22)
  (p6 ph24 d17 p7 ph25 d17 p6 ph24)
  (p6 ph24 d17 p7 ph25 d17 p6 ph24)
  (p6 ph22 d17 p7 ph23 d17 p6 ph22)
  (p6 ph24 d17 p7 ph25 d17 p6 ph24)
  (p6 ph24 d17 p7 ph25 d17 p6 ph24)
  (p6 ph22 d17 p7 ph23 d17 p6 ph22)
  (p6 ph22 d17 p7 ph23 d17 p6 ph22)
  (p6 ph24 d17 p7 ph25 d17 p6 ph24)
  (p6 ph22 d17 p7 ph23 d17 p6 ph22)
  (p6 ph22 d17 p7 ph23 d17 p6 ph22)
  (p6 ph24 d17 p7 ph25 d17 p6 ph24)
  (p6 ph22 d17 p7 ph23 d17 p6 ph22)
  (p6 ph22 d17 p7 ph23 d17 p6 ph22)
  (p6 ph24 d17 p7 ph25 d17 p6 ph24)
  (p6 ph24 d17 p7 ph25 d17 p6 ph24)
  (p5 ph23)
  lo to 4 times l1
  (p17 ph26)
  go=2 ph31
  d11 wr #0 if #0 id0 ip1 zd
  lo to 3 times td1
exit
```

```
ph1 =0 2 2 0 1 3 3 1
ph22=0 2 0 2 1 3 1 3
ph23=1 3 1 3 2 0 2 0
ph24=2 0 2 0 3 1 3 1
ph25=3 1 3 1 0 2 0 2
ph26=1 3 1 3 2 0 2 0
ph31=0 2 2 0 1 3 3 1
ph29=0

;hl2: ecoupler low power level for presat
;hl3: ecoupler power level for 1st pulse
;         and MLEV spinlock (7 dB)
;
;p1 : 90 degree ecoupler pulse
;p5 : 60 degree ecoupler pulse
;p6 : 90 degree ecoupler pulse (10-15usec)
;p7 :180 degree ecoupler pulse
;p17: trim pulse (2.5msec)
;p18: presaturation during recycle delay
;
;d0 : incremented delay (2D) (3usec)
;d17: compensation time during mixing
(10-40usec)
;d11: delay for disk I/O (30 msec)
;d12: delay for power switching (20 usec)
;d13: short delay (e.g. to compensate delay
line) (3 usec)
;
;L1 : loop for MLEV cyle:
;τm=((p6*64)+p5+(32*d17))*L1+(p17*2)
;
;in0: 1/(2*SW) = DW for AQ_mode=qseq
;nd0: 2
;NS : 8 * n
;DS : 2 or 4
;td1: number of experiments
;MC2: TPPI
```

```
;noesyprtp.m
;noesy
;with presat during RD and mixing time
;quad in t1 using TPPI

1 ze
2 d11
3 d12 hl2
  d6
  p18 ph29
  d13
  d12 hl1
  p1 ph1
  d0
  p1 ph2
  d9 hl2
  p19 ph29
  d13
  d12 hl1
  p1 ph3
  go=2 ph31
  d11 wr #0 if #0 id0 ip1 zd
  d11 id9
  lo to 3 times td1
exit

ph1= 0 2
ph2= 0 0 0 0 0 0 0 0
     2 2 2 2 2 2 2 2
ph3= 0 0 2 2 1 1 3 3
ph31=0 2 2 0 1 3 3 1
     2 0 0 2 3 1 1 3
ph29=0

;DP=hl1: ecoupler high power level
;hl2: ecoupler power level for presaturation
;p1 :  90 degree transmitter high power
;p18: presaturation during relaxation delay
;p19: presaturation during mixing time-
NOTE tm=p19+d9
;d0 : incremented delay  (2D)        [3 usec]
;d6 : additional relaxation delay
;d9 : set to 5ms
;d11: delay for disk I/O             [30 msec]
;d12: delay for power switching      [20 usec]
;d13: short delay                    [3 usec]
;NS : 8 * n
;DS : 2 or 4
;MC2: TPPI
;in0: 1/(2*swh)=dw
```

```
;inv4jr1d.bv
;H-1/X correlation via heteronuclear zero
;and double quantum coherence 1-D
;phase sensitive using TPPI
;with 1-1JR excitation and
;refocusing pulses
;with decoupling during acquisition
;Sklenar, V. and Bax, A., J. Mag. Res. 74,
469-479 (1987)

;d6 : delay for refocusing pulse water ;
        suppression (2*d5)
;d11: delay for disk I/O              [30 msec]
;d12: delay for power switching   [20 usec]
;d13: short delay                    [3 usec]
;NS: 16 * n
;DS: 2 or 4
;MC2: TPPI
;cpd: cpd-decoupling according to sequence
;defined by cpdprgb

1 ze
2 d11
3 d1 dbhi
 d13
 d12 hl1
 p1 ph1
 d5
 p0 ph5
 d2 dbhi
 (p3 ph2):f3
 d13
 (p3 ph4):f3
 d13
 p1 ph3
 d6
 p0 ph6
 d2 dblo
 go=2 ph31 cpdb
 d11 wr #0
exit


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


;hl1: ecoupler high power level
;dl0: power level for dlo mode
;p0 :  90 degree transmitter for 11JR
;p1 :  90 degree transmitter high power
pulse
;p3 :  90 degree decoupler high power pulse
;p31:  90 degree pulse for cpd decoupling
;d1 : relaxation delay
;d2 : 1/(2J)XH
;d5 : delay for excitation pulse water ;
        suppresion
```

```
;hmqcjrY.bv
;2D H-X correlation via heteronuclear zero
and double quantum
;coherence
;phase sensitive using States-TPPI
;with 11-JR water suppression
;and decoupling during acquisition
;X-nucleus on Y channel
;BFV 2/16/94

1 ze
2 d11
3 d1 dbo
4 d13
 d12 hl1
 p1 ph1
 d5
 p0 ph5
 d2 dbhi
 p3:f3 ph2
 d0
 p1 ph3
 d6
 p0 ph6
 d0
 p3:f3 ph4
 d2 dblo
 go=2 ph31 cpdb
 d11 wr #0 if #0 ip2 zd
 d11 dbo
 lo to 3 times 2
 d11 ip31 id0
 d11 ip31
 lo to 2 times l0
exit

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

;hl1: ecoupler high power level
;dbl0: power level for cpd decoupling
;p0 : transmitter 90 degree high power
pulse
;p1 : transmitter 90 degree high power
pulse
;p3 : decoupler 90 degree high power pulse
;p31: decoupler 90 degree low power pulse
;

;d0 : incremented delay
;d1 : relaxation delay
;d2 : 1/(2J)HX
;d5 : delay for binomial water suppression
;d6 : delay for binomial water suppression
(2*d5)
;d11: delay for disk I/O
;d12: delay for power switching
;d13: short delay
;l0 : number of complex blocks
;in0:1/2*DW(X)
;nd0: 4
;NS : 16*n
;DS : 16
;td1: number of experiments
;MC2: States-TPPI
;cpd: cpd-decoupling according to sequence
defined by cpdprg
```

```
;fasthsqc.al
;HSQC with no phase cycling,
;TPPI-STATES in t1
;D. Marion et al., JMR 85, pp. 393-399
;(1989)
;X nucleus on Y channel
;Solvent suppr. with scrabling
;pulses and presat.


d6=3u
d8=aq
d11=50m
d12=5m
d25=p3-p1
p5=p1*(8/3)
p8=p1*111 ;10,000 deg non-selec. 1H pul
p9=p1*55 ;5,000 deg non-selec. 1H pulse
p24=d4 ;

;HL2=presat power
;p18=presat pulse
;d4=1/4J


1 ze
2 p8 ph1
  p9 ph2
3 d12
  d12
  d12
4 10u hl2 dbo
  p18 ph0
  10u hl1
  p1 ph1
  d4 dbhi
  (d25 p1*2 ph1) (p3*2 ph1):f3
  d4
  p22 ph1          ;Messerle pulse = 1.8 msec
  (d25 p1 ph2) (p3 ph3):f3  ;1H 90, X 90
  d0
  (p1 ph1 d6 p5 ph2 d6 p1 ph1) ;180 1H
  d0
  (d25 p1 ph1) (p3 ph1):f3
  d4
  (d25 p1*2 ph1) (p3*2 ph1):f3
  p24:e dblo
  d7                                ;DE
  d6 adc
  d8 cpdb
  rcyc=2 ph31               ;set ns=1
d11 dbo wr #0 if #0 zd    ;write to disk
d12 ip3
  p8 ph1
  p9 ph2
lo to 3 times 2
d12 ip31
d12 ip31
d12 id0
lo to 4 times l3
exit

  ph0=1
  ph1=0
  ph2=1
  ph3=0 2        ;cycle if ns > 1.  otherwise
                 ;set ph3=0
  ph31=0 2
```

```
;hohmY3d.bv
;derived from invtoc3d.gab and tocinv3d.al
;3D with isotr. mix via "clean" MLEV17
;H-X correlation via HMQC
;X-nucleus on Y channel
;heteronuclear decoupling during acq
;one power level for excitation and SL
;quadrature in t1 and t2 by States-TPPI
;Griesinger et al., JACS 110, 7870 (1988)
;Wijmenga et al., JMR 84, 634 (1989)
;Edison et al., JMR 92, 434 (1991)

1 d11 ze
2 d11 dbo
3 d11*2
4 do hl2
  d12 dbhi
  p18 ph29
  d12 hl3
  p1 ph1
  d0
  p4:f3 ph2
  d0
  (p17 ph26)
5 (p6 ph22 d6 p7 ph23 d6 p6 ph22)
  (p6 ph24 d6 p7 ph25 d6 p6 ph24)
  (p6 ph24 d6 p7 ph25 d6 p6 ph24)
  (p6 ph22 d6 p7 ph23 d6 p6 ph22)
  (p6 ph24 d6 p7 ph25 d6 p6 ph24)
  (p6 ph24 d6 p7 ph25 d6 p6 ph24)
  (p6 ph22 d6 p7 ph23 d6 p6 ph22)
  (p6 ph22 d6 p7 ph23 d6 p6 ph22)
  (p6 ph24 d6 p7 ph25 d6 p6 ph24)
  (p6 ph22 d6 p7 ph23 d6 p6 ph22)
  (p6 ph22 d6 p7 ph23 d6 p6 ph22)
  (p6 ph24 d6 p7 ph25 d6 p6 ph24)
  (p6 ph22 d6 p7 ph23 d6 p6 ph22)
  (p6 ph22 d6 p7 ph23 d6 p6 ph22)
  (p6 ph24 d6 p7 ph25 d6 p6 ph24)
  (p6 ph24 d6 p7 ph25 d6 p6 ph24)
  (p5 ph23)
  lo to 5 times l1
  (p17 ph26)
  d2
  p3:f3 ph3
  d5
  p2 ph4
  d5
  p3:f3 ph5
  d2 dblo
  go=4 ph31 cpdb
  d11 wr #0 if #0 ip1 zd
  lo to 3 times 2
  d11 ip31 id0
  d11 ip31
  lo to 3 times l0
  d11 rd0 ip3
  d11 rp1
  lo to 4 times 2
  d11 ip31 id5
  d11 ip31
  lo to 4 times l5
exit

ph1=0 0 0 0 2 2 2 2
ph2=0
ph3=0 2
ph4=0
ph5=0 0 2 2
ph22=0
ph23=1
ph24=2
ph25=3
ph26=1
ph29=0
ph31=0 2 2 0 2 0 0 2

;hl1: ecoupler high power level
;hl2: ecoupler power level for solvent presat
;hl3: ecoupler low power level for spin-lock
;p1 :  90 degree transmitter high power
;p2 : 180 degree transmitter high power
;p3 :  90 degree decoupler high power pulse
;p4 : 180 degree decoupler high power
;p5 :  60 degree transmitter low power
;p6 :  90 degree transmitter low power
;p7 : 180 degree transmitter low power
;p17: trim pulse                [1-2 ms]
;p18: relaxation delay with presaturation
;p31: cpd 90 degree low power dec pulse
;d0 : incremented delay (F1 in 3D)  [3 usec]
;d2 : 1/((2J)XH)          [3 msec]
;d6 : clean TOCSY delay, d6=p6
(d6=2.5*p6  -FFD)
;d5: incremented delay (F2 in 3D)   [3 usec]
;d11: delay for disk I/O          [30 msec]
;d12: delay for power switching  [200 usec]
;in0: 1/(4 * SW(H)) = (1/2) DW(H) for t1
;nd0: 4
;iN5: 1/(4 * SW(X)) = (1/2) DW(X) for t2
;NS: 8 * n
;DS: 16
;l0: number of experiments in F1
;l1: loop for MLEV cycle - ((p6 *
96)+p5)*l1)+p17*2 = mixing time
;l5: number of experiments in F2
;MC2: TPPI in F1
;MC2: TPPI in F2
```

```
;hohaha15n.bv
;3D HOHAHA-HSQC sequence
;with DIPSI isotrop. mixing
;and Messerle spin-lock for
;solvent suppression
;derived from hohaha15n.jgp
;and hohaha15n.sg
; NOTE!! CHOICE OF PRESAT
;OR HOMOSPOIL SUPPRESSION
;DURING NOE COMP. DELAY -
;COMMENT OUT/IN APPROPRIATE
;LINES (3 FOR PRESAT, 2 FOR HS)
;Added 15N dec. 180 in t1
;and used -90,180 phasing
;in t1 and t2
;BFV 10/14/94
;References:
;*Spinlock solv. suppress.*
;Messerle, B. A., et al,
;1989, JMR 85, 608-613.
;*Samp. delay -90,180 trick*
;Bax, A., et al, 1991
;JMR 91, 174-178

;**** pulses ****
;p1=proton 90 at hl1
;p3=nitrogen 90 at dbhi, yl=3dB
p4=2.25m  ; 1/4J
;p5 1H 90 at hl3 (ca. 15 dB)
p9=700u        ;SL-pulse ;DIPSI I
p10=p5*320.0/90.0 ;pulses V
p11=p5*410.0/90.0
p12=p5*290.0/90.0
p13=p5*285.0/90.0
p14=p5*30.0/90.0
p15=p5*245.0/90.0
p16=p5*375.0/90.0
p17=p5*265.0/90.0
p18=p5*370.0/90.0
p19=p5*60.0/90.0  ;     _I
;p21=presat pulse = ca. 1s
p22=2.0m; Messerle pulse
;p31=nitrogen 90 at dbl0 ca. 150us
;****  delays, increments  ****
;d0=(in0/2)-p3-(p1/3.14159)
d1=1.0m
;d3 = half mixing time
d4=2.25m
;d7 ca 100u = 1.6*dw
d8=aq
d10=(in10/2)-p1-(p3*2/3.14159)
d11=50m
d12=11m
d26=p3-p1

;in0= proton incr = 1/(2*sw)
;in10=nitrogen incr = 1/(2*sw)
;****  power levels  ****
;hl1 = 3dB
;hl2 = ca. 50dB
;hl3 = 11dB
;dbl0 = 14.7dB
;yl  = 2dB
;****  freq. settings  ****
;f1 = hydrogen frequency on water
;f3 = nitrogen frequency

**** start pulse program ***
        ze
        d11
2       d11
      d12*2
3     d12*3
4     d12*2
5     d12*3

;*******presaturation*******
6       d1 hl2 dbhi
        p21 ph0
        5u
        20u hl1

;*******start 90-degree *******
        p1 ph2

;*******proton evolution*******
        d0      ;t1/2
        (p3*2 ph0):f3
        d0      ;t1/2

;*******SL pulse*******
        p9 ph1
        2u
        2u hl3

;*******dipsi-2 mixing*******
7       (p10 ph20 p11 ph19 p12 ph20
              p13 ph19 p14 ph20)
        (p15 ph19 p16 ph20 p17 ph19
              p18 ph20)
        (p10 ph19 p11 ph20 p12 ph19
              p13 ph20 p14 ph19)
        (p15 ph20 p16 ph19 p17 ph20
              p18 ph19)
        (p10 ph19 p11 ph20 p12 ph19
              p13 ph20 p14 ph19)
        (p15 ph20 p16 ph19 p17 ph20
              p18 ph19)
        (p10 ph20 p11 ph19 p12 ph20
```

```
                p13 ph19 p14 ph20)
        (p15 ph19 p16 ph20 p17 ph19
                p18 ph20)
;       (p19 ph21)
        lo to 7 times l2

;******total duration=115.11 * p5 * l2***
        2u
        2u hl1

;*******SL pulse*******
;       (p9 ph1)
        2u

;*******read pulse*******
        (p1 ph0)
;       10u hl2
;       p20 ph0 ;presat
;       10u hl1
        p8:h    ;homospoil
        d3      ;half mixing time
        (p1 ph0)
        d4
        (d26 p1*2 ph4) (p3*2 ph10):f3
        d4

;*******water trim pulse*******
        (p22 ph0)
        2u
        (p1 ph1) (p3 ph11):f3

;*******N15 evolution*******
        d10     ;t2/2
        (p1*2 ph0)
        d10     ;t2/2
        (p1 ph0) (p3 ph10):f3
        d4
        (d26 p1*2 ph3) (p3*2 ph10):f3
        2u
        (2u ph0)
        d7 dblo
        p4:e

;******acquire FID*****
        3u adc
        d8 cpdb
dbo
rcyc=2 ph31
d11 wr #0 if #0 zd
d12 ip2
d12 ip2
lo to 3 times 2
d12 id0
d12 ip31
d12 ip31
lo to 4 times 128
d12 ip11
d12 rd0
lo to 5 times 2
d12 id10
d12 ip31
d12 ip31
lo to 6 times 26
dbo
exit

ph0=0
ph1=1
ph2=(8) 1 5
ph3=0 0 1 1 2 2 3 3
ph4=0 0 2 2 1 1 3 3
ph10=0
ph11=0 0 0 0 0 0 0 0
       2 2 2 2 2 2 2 2
ph19=(3600)0
ph20=(3600)1800
ph21=(3600)900
ph31=0 2 2 0 2 0 0 2
       2 0 0 2 0 2 2 0
```

```
;noei4Ypr3d.bv
;3D sequence with homonuclear
;correlation via dipolar coupling
;dipolar coupling may be due to noe
;or chemical exchange.
;H-1/X correlation via heteronuclear zero
;and double quantum coherence.
;X-nucleus on Y channel
;phase sensitive using TPPI-States
;with decoupling during acquisition
;with presaturation during relaxation delay
and mixing time


1 d11 ze
2 d11 dbo
3 d11*2
4 dbo hl2
 d12 dbhi
 p18 ph29
 d13
 d12 hl1
 p1 ph1
 d0
 p4:f3 ph7
 d0
 p1 ph2
 d13
 d12 hl2
 p19 ph29
 d13
 d12 hl1
 p1 ph3
 d2
 p3:f3 ph5
 d5
 p2 ph4
 d5
 p3:f3 ph6
 d2 dblo
 go=4 ph31 cpdb
 100m dbo wr #0 if #0 ip1 zd
 lo to 3 times 2
 d11 ip31
 d11 ip31 id0
 lo to 3 times l0
 d11 rd0 ip5
 d11 rp1
 lo to 4 times 2
 d11 ip31
 d11 ip31 id5
 lo to 4 times l5
exit

ph1=0 0 0 0 2 2 2 2
ph2=0
ph3=0
ph4=0
ph5=0 2
ph6=0 0 2 2
ph7=0
ph29=0
ph31=0 2 2 0 2 0 0 2


;hl1: ecoupler high power level
;hl2: ecoupler power level for presaturation
;dbl0: power level for dlo mode (17dB)
;p1 :  90 degree transmitter high power
;p2 : 180 degree transmitter high power
;p3 :  90 degree decoupler high power pulse
;p4 : 180 degree decoupler high power
;p18: presaturation during relaxation delay
;p19: presaturation during mixing time
;p31: 90 deg low power X pulse for slave
        timer (cpd-sequence)
;d0 : incremented delay (t1 in 3D)  [3 usec]
;d2 : 1/((2J)XH)
;d5: incremented delay (t2 in 3D)   [3 usec]
;d11: delay for disk I/O    [30 msec]
;d12: delay for power switching    [20 usec]
;d13: short delay (e.g. to compensate delay
        line)    [3 usec]
;in0: 1/(4 * SW(H)) = (1/2) DW(H)
;nd0: 4
;in5: 1/(4 * SW(X)) = (1/2) DW(X)
;NS: 4 * n
;DS: 32
;l0: number of complex blocks in F1
;l5: number of complex blocks in F2
;MC2: States-TPPI in F1
;MC2: States-TPPI in F2
;cpd: cpd-decoupling according to sequence
        defined by cpdprg
```

```
;cbcaconh.sg
#undef COMMENT
#ifdef COMMENT
;sg 1/28/92

;IN0 = IN0
;LOOP1 = (d18-IN0)%in0 + 1 = d18%IN0
;LOOP1+LOOP2 = total complex 13C incs
= d17%IN0

#endif

#define        IN0     59.2u
#define LOOP1          36
#define LOOP2          14

#undef CARBON
#define CARBON
#undef NITROGEN
#define NITROGEN

;t=f1 = hydrogen frequency on water
;p1=proton 90 at hl1
;p30=proton 90 at hl3
;hl1 = 1dB
;hl3 = dipsi decoupling power 15db
;f2 = carbon frequency switched from
        46ppm to 177ppm
;p4=low power carbon 90 58.8us at dl0 at
        500 MHz
;p6=low power carbon 90 26.3us at dl2
;p20=low power carbon 90 63.9us at dl4
;p21=low power carbon 90 28.6u at dl5
p11=202.4u  ;shaped offset sinc
p22=1.8m       ;water trim pulse
p23=0.1m       ;water trim pulse
;p24     ;opening the receiver
;f3 = nitrogen frequency
;p7 = high power nitrogen 90 42us
;p31 = low power pulse 160 us at dbl0
;p18=presat pulse = ca. 1s

p2=2.25m
d0=4u
;d7 ca 100u = dw
d5=5m          ;universal dd, id, ip delay
d8=aq
d11=50m
d13=217.6u  ; freq switching 4/delta freq
d14=d13-p21*2.0
d15=5.4m
d16=4u
d17=3.1m; 7.1m; 10.5m; 3.1m  ;cb to ca
                transfer time/2
d18=2.1m+IN0
d19=d17-d18+IN0
d26=d0+p11+4u+d18-IN0
d27=4u
d28=d0+IN0*LOOP1+d17-d26-d27
d10=11m+d16+p21*2.0+10u
d20=11m-d15
d21=p4-p1
d22=p4-p7
d23=p7-p1
d24=p20-p7
d25=p6-p1
d30=d17+d0
d29=d30-IN0*LOOP1
;in0=in30=in29=in18=in19=in27
;in10=in20=in15=in16

        ze
2       d5
3       d5 dd18
4       2u
5       2u
;*******presaturation*******
6       50m hl2
        50u dbhi
        1m dlo dl2
        50u o1
        50u o2
        d1
        p18 ph0
        10u hl1
;*******start 90-degree on ha/hb*******
        (p1 ph2)
        1.5m
        2u
        (d25 p1*2 ph0) (p6*2 ph7):f2
        2u
        1.5m dl0
        (p22 ph2)       ;water trim pulse
        2u
;*******     inept to ca/cb      *******
        (p1 ph7) (p4 ph1):f2
        4u
        d0 hl3
        (p11 ph10):dp1 ;shaped carbonyl
        4u
        d18 dl2
        ;d18+d19=d17, d0+d18 = 2.1m
        d19 cpdts
        (p6*2 ph17):f2
        d30
        (p11 ph10):dp1  ; Bloch-Siegert
        4u
        4u dl0
```

```
;******* inept to ca *******
        (p4 ph19):f2
        4u
        (p11 ph10):dp1  ;shaped carbonyl
        10u
        3.5m dl2        ;
        (p6*2 ph9):f2
        4u
        (p11 ph10):dp1;shaped carbonyl
        3.5m dl0
        10u
;*******inept to carbonyl      *******
        (p4 ph0):f2
        50u o2; shifting frequency to C'
        10u dl4
        (p20 ph8):f2
        1m
        3.5m dl5
        d13 o2;          shifting frequency to
                ; Cα    -18248 Hz
        (p21*2 ph10):f2
        d14
        5.5m o2   ;shifting frequency  to C'
        1m dl4
        (p20*2 ph10):f2(d24 p7*2 ph10):f3
        11m
        d13*2
;*******  inept to nitrogen    *******
        (p20 ph18):f2
        4u to
        4u o1   ;shifting frequency to mid of
                amid protons
        0.1m cpdts
        (p7 ph9):f3     ;inept to nitrogen
        d10
        (p20*2 ph10):f2 (d24 p7*2 ph4):f3
        d20 dl5
        d15 to
        10u o2          ;shifting frequency
                to cα -18248
        (p21*2 ph10):f2
        d16 hl1 o1      ;shifting frequency
                back to water
;*******inept to protons       *******
        (d23 p1 ph0) (p7 ph0):f3
        2u
;     (p23 ph2)          ;water trim pulse
        2.25m
        (d23 p1*2 ph0) (p7*2 ph10):f3
        2.25m
;     (p23*2 ph0)      ;water trim pulse
        (2u ph0)
        d7 dblo
        p24:e
        3u adc
        d8 cpdb
        dbo
        rcyc=6 ph31
        d11 wr #0 if #0 zd

#ifdef CARBON
d5 ip1
lo to 4 times 2
d5 id0
d5 dd30
d5 id19
d5 ip31
d5 ip31
lo to 3 times LOOP1

;*******here starts LOOP2*******
12      d5
        d5
13      20m
14      10m
15      30m
;*******presaturation*******
16      50m hl2
        50u dbhi
        1m dlo dl2
        50u o1
        50u o2
        d1
        p18 ph0
        10u hl1
;*******start 90-degree on ha/hb*******
        (p1 ph2)
        1.5m
        2u
        (d25 p1*2 ph0) (p6*2 ph7):f2
        2u
        1.5m dl0
        (p22 ph2)       ;water trim pulse
        2u
;*******       inept to ca/cb   *******
        (p1 ph7) (p4 ph1):f2
        4u
        d26 hl3
        d27 cpdts
        (p11 ph10):dp1 ;shaped pulse on C'
        4u
        d28 dl2
        (p6*2 ph17):f2
        d29
        (p11 ph10):dp1;shaped pulse on
                ;carbonyl Bloch-Siegert
        4u
```

```
        4u dl0
;*******inept to ca     *******
        (p4 ph19):f2
        4u
        (p11 ph10):dp1 ;shaped pulse on C'
        10u
        3.5m dl2        ;be careful when
                        ;using shaped pulses
        (p6*2 ph9):f2
        4u
        (p11 ph10):dp1;shaped pulse on C'
        3.5m dl0
        10u
;*******   inept to carbonyl  *******
        (p4 ph0):f2
        50u o2; shifting frequency to C'
        10u dl4
        (p20 ph8):f2
        1m
        3.5m dl5
        d13 o2          ;shifting frequency
                        ;to Ca -18248 Hz
        (p21*2 ph10):f2
        d14
        5.5m o2         ;shifting frequency
                        ;back to carbonyl
        1m dl4
        (p20*2 ph10):f2(d24 p7*2 ph10):f3
        11m
        d13*2
;*******  inept to nitrogen    *******
        (p20 ph18):f2
        10u to
        10u o1          ;shifting frequency
                ;to mid of amid protons
        0.1m cpdts
        (p7 ph9):f3     ;inept to nitrogen
        d10
        (p20*2 ph10):f2 (d24 p7*2 ph4):f3
        d20 dl5
        d15 to
        10u o2  ;shifting freq to ca -18248
        (p21*2 ph10):f2
        d16 hl1 o1      ;shifting frequency
                        ;back to water
;*******  inept to protons     *******
        (d23 p1 ph0) (p7 ph0):f3
        2u
;       (p23 ph2)       ;water trim pulse
        2.25m
        (d23 p1*2 ph0) (p7*2 ph10):f3
        2.25m
;       (p23*2 ph0)     ;water trim pulse
        (2u ph0)
        d7 dblo
        p24:e
        3u adc
        d8 cpdb
dbo
rcyc=16 ph31
d11 wr #0 if #0 zd
d5 ip1
lo to 13 times 2
d5 id27
d5 dd29
d5 ip31
d5 ip31
lo to 14 times LOOP2
#endif

#ifdef NITROGEN
d5 ip9
d5 rd0
d5 rd30
d5 rd29
d5 rd27
d5 rd18
d5 rd19
lo to 2 times 2
d5 dd10
d5 dd15
d5 id16
d5 id20
d5 ip31
d5 ip31
lo to 2 times 32
#endif
do
exit

ph0=0
ph1=0
ph2=1
ph4=0 0 0 0 0 0 0 0
    2 2 2 2 2 2 2 2
ph7=0 2
ph8=0 0 2 2
ph9=0 0 0 0 2 2 2 2
ph17=0 0 0 0 0 0 0 0
     1 1 1 1 1 1 1 1
     2 2 2 2 2 2 2 2
     3 3 3 3 3 3 3 3
ph18=(360)51            ;adjusted for BS
ph19=0
ph31=0 2 2 0 2 0 0 2
     2 0 0 2 0 2 2 0
ph10=0
ph20=0
```

```
;ct-hsqced31.jgp  for filtering arom. or C'
;original expt
;2D 1H-13C hsqc with BB dec of 13C
;Grzesiek and Bax (1993) J. Biol. NMR 3,
;pp.185-204.
;constant time 2T tuned to 1/Jcc
;need extra 13C ext. amp !!must use dlo !!
;p3 ca 18-20us 13C90 at dl0
;p13 26.3us 90 13C at dl1 null on carbonyl
;            when carrier set to 43 ppm
;p31 ca 60-65 us at dl2 for cpd decoupling
;p6 dp0 sinc1.0 180 13C' ( 240 us ) with
;dp0 40db approx 20 kHz offset
;set d7 to get 0 linear phase corr. 1.6*DW
;states with messerle water trim pulse
d26=p3-p1
d2=1.7m
d8=aq
p11=1m
d12=7m
d20=10u
d21=13.3m
d14=13.3m
p22=1.7m
d0=10u
d15=3m
d16=10.3m


1 ze
2 d12 dlo
3 d12*4
4 d12*3
5 d1 hl2
 p18 ph0       ;on resonance presat at Hl2
 10u hl1         ;set to HP 1H
 p1 ph0          ;1H 90
 d2 dl0          ;1/4JXH 1.7ms
 2u
 (d26 p1*2 ph0) (p3*2 ph1):d   ;1H,X 180
 2u
 d2
 (p11 ph0)      ;Messerele trim pulse < 1ms
 2u
 (d26 p1 ph1) (p3 ph2):d     ;1H,X 90
 d0                  ;t1/2
 (p1*2 ph0)            ;180 1H,
 (p6 ph0):dp0         ;180 C' shaped pulse
 10u
 d14 dl1               ;time T
 (p13*2 ph6):d       ;apply selective X 180
 d21                   ;T-t1/2
 10u dl0               ;set low power X
 (p6 ph0):dp0     ;apply shaped at pwr dp0
 (p1*2):c8          ;1H 180 compensation
 10u dl0
 (d26 p1 ph0) (p3 ph3):d      ;90 1H, 90 X
 d2
 (d26 p1*2 ph4) (p3*2 ph5):d ;180 1H&X
 d7          ;adjust for flat baseline 1.6*DW
 2u ph0                ;set rec phase
 p22:e dl2          ;open rec, wait period d2
 10u adc cpds            ;start adc and cpd
 d8                   ;set to acq time
 do
 rcyc=2 ph31
 d11 wr #0 if #0 zd   ;write + increment file
 d12 ip2            ;increment ph2 for States
lo to 3 times 2
 d12 ip31     ;increment rec for States TPPI
 d12 ip31
 d12 id0              ;increment d0 by in0
 d12 dd21           ;decrement d21 by in21
lo to 4 times l3
exit
ph0=0
ph1=1 3
ph2=0
ph3=0 0 0 0 0 0 0 0
   2 2 2 2 2 2 2 2
ph4=0
ph5=0
ph6=(360) 83 83 263 263 173 173 353 353
;adjust for low power
ph31=0 2 0 2 2 0 2 0
    2 0 2 0 0 2 0 2
;hl1 ecoupler high power
; p1 90 1H high power at hl1
; d1 relax delay
; d2 1/4JXH
; in0 dw/2 X dimension
; in21 = in0
; ns = 16*n
; ds = 16
```

```
;cthsqcRNA.jgp
;2D 1H-13C hsqc with broad-band
;decoupling of 13C
;Vuister & Bax (1992) JMR 98, pp 428
;Santoro + King (1992) JMR 97, 202-207
;constant time 2T tuned to 1/Jcc
;states quad in t1 with messerle water trim
;pulse
;15N decoupling removed 11/4/93 jgp
d26=p3-p1
d8=aq
p22=2m
d12=11m
d11=50m
d0=3u
p9=10u
d5=(p1*2)+(d0*2)

1 ze
2 d12
3 d12*4
4 d12*3
5 d1 hl2          ;recycle delay - approx
1.5s
;  p18 ph0        ;on resonance presat at
Hl2
 10u hl1          ;set to Hi Power 1H
 p1 ph0           ;1H 90
 d2 dhi           ;1/4JXH 1.4ms  - 13C hi
pwr
 2u
 (d26 p1*2 ph0) (p3*2 ph1):d      ;1H,X
180
 2u
 d2
 (p22 ph0)                ;Messerele trim
pulse  2ms
 2u
 (d26 p1 ph1) (p3 ph2):d     ;1H,X 90
 d0                    ;t1/2
 (p1*2 ph0)                ;180 1H
 d0                     ;t1/2
 d21                    ;T-t1/2
 (p3 ph6 2u p3*2 ph7 2u p3 ph6):d   ;13C
180 , composite
 d21                    ;T-t1/2
 d5                     ; compensate for phase
 (d26 p1 ph0) (p3 ph3):d     ;90 1H, 90 X
 d2
 (d26 p1*2 ph4) (p3*2 ph5):d  ;180 1H,
180 C
 d7                     ;adjust for flat baseline
1.6*DW
 p9:e ph0                ;set rec phase
 d2 dlo                  ;open rec, wait
period d2
 10u adc cpds              ;start adc and cpd
 d8                     ;set to acq time
 do
 rcyc=2 ph31
 d11 do wr #0 if #0 zd       ;write +
increment file
 d12 ip2                 ;increment ph2 for
States
lo to 3 times 2
 d12 ip31                ;increment rec for
States TPPI
 d12 ip31
 d12 id0                 ;increment d0 by in0
 d12 dd21                 ;decrement d21 by
in21
lo to 4 times l3
exit
ph0=0
ph1=1 3
ph2=0
ph3=0 0 0 0 0 0 0 0
   2 2 2 2 2 2 2 2
ph4=0
ph5=0
ph6=1 1 3 3 2 2 0 0
ph7=2 2 0 0 3 3 1 1
ph31=0 2 0 2 2 0 2 0
    2 0 2 0 0 2 0 2
;hl1 ecoupler high power
; p1 90 1H high power at hl1
; p3 13C 90 high power - approx 16us
; p31 13C low power for Garp decoupling
approx 60us
; d1 relax delay
; d2 1/4JXH approx 1.4ms
; in0 dw/2 X dimension
; in21 = in0
; ns = 16*n
; ds = 4
; d12 =7ms switch
; p18 - for presat if needed at hl2 -
; p22 spinlock trim pulse for h2o
suppression <2ms
; d0 - inital t1 period 3us
; d21 1/Jcc or 2/Jcc - either 9ms or 18ms
; d7 1.6*dw
; p9 2us
; d8 set to acq time
;
```

```
;hcchtoc.al
;3D 13C HCCH-TOCSY
;sg 10-15-91
;t=f1=hydrogen frequency on water
;p1= proton 90 at hl1
;p18 = proton presaturation at hl2
;d1 = relaxation delay
;
;f2 = carbon frequency at 46 ppm
;p3 = high power carbon 90 ca. 10us
;p4 = shaped pulse on carbonyl 180 degree
p5 = 2m          ;low power trim pulse at dl0
;p6 low power carbon 90 30-35 us for
dipsi/flopsy at dl0
p10=p6*.55555
;p31 low power carbon ca. 50us 90 for
waltz decoupling dl1
d21=p3-p1
d6=p3*4+d0*2+4u
;d7=de
d8=aq
d12=p4 + d10*2 +p3*0.6366 +p1*2
p7=1.5m
p8=p3+1.5u
p9=p3-1.5u

1       ze
2       50m
3       10m
4       10m
5       15m
6       10u hl2
        d1 o1
        p18 ph0
        10u o1
        10u hl1
;****************************
        p1 ph1          ;starting 90
        1.5m dhi
        d0
        (p3 ph2):f2   ;composite 180 13C
     2u
        (p3*2 ph13):f2
        2u
        (p3 ph2):f2
        d0
        p1*2 ph3
        1.5m
        d6      ;compensation delay
;*********************************
        (d21 p1 ph4) (p3 ph5):f2       ;first
inept
        2u
        d10 dlo
        (p4 ph10):dp1          ;Carbonyl
180
        4u
        1m
        (p1*2 ph6)
        2u
        d10 dhi
        (p3*2 ph7):f2
        2u
        0.5m dlo
        0.5m dl0
        6u
        d12
;*********************************
        (p5 ph12):f2    ;SL-pulse
;*******       start DIPSIy*********
8       (p10*4.9 ph15):f2
        (p10*7.9 ph16):f2
        (p10*5.0 ph15):f2
        (p10*5.5 ph16):f2
        (p10*0.6 ph15):f2
        (p10*4.6 ph16):f2
        (p10*7.2 ph15):f2
        (p10*4.9 ph16):f2
        (p10*7.4 ph15):f2
        (p10*6.8 ph16):f2
        (p10*7.0 ph15):f2
        (p10*5.2 ph16):f2
        (p10*5.4 ph15):f2
        (p10*0.6 ph16):f2
        (p10*4.5 ph15):f2
        (p10*7.3 ph16):f2
        (p10*5.1 ph15):f2
        (p10*7.9 ph16):f2
        (p10*4.9 ph15):f2
        (p10*7.9 ph16):f2
        (p10*5.0 ph15):f2
        (p10*5.5 ph16):f2
        (p10*0.6 ph15):f2
        (p10*4.6 ph16):f2
        (p10*7.2 ph15):f2
        (p10*4.9 ph16):f2
        (p10*7.4 ph15):f2
        (p10*6.8 ph16):f2
        (p10*7.0 ph15):f2
        (p10*5.2 ph16):f2
        (p10*5.4 ph15):f2
        (p10*0.6 ph16):f2
        (p10*4.5 ph15):f2
        (p10*7.3 ph16):f2
        (p10*5.1 ph15):f2
        (p10*7.9 ph16):f2
;
        (p10*4.9 ph16):f2
```

```
(p10*7.9 ph15):f2
(p10*5.0 ph16):f2
(p10*5.5 ph15):f2
(p10*0.6 ph16):f2
(p10*4.6 ph15):f2
(p10*7.2 ph16):f2
(p10*4.9 ph15):f2
(p10*7.4 ph16):f2
(p10*6.8 ph15):f2
(p10*7.0 ph16):f2
(p10*5.2 ph15):f2
(p10*5.4 ph16):f2
(p10*0.6 ph15):f2
(p10*4.5 ph16):f2
(p10*7.3 ph15):f2
(p10*5.1 ph16):f2
(p10*7.9 ph15):f2
(p10*4.9 ph16):f2
(p10*7.9 ph15):f2
(p10*5.0 ph16):f2
(p10*5.5 ph15):f2
(p10*0.6 ph16):f2
(p10*4.6 ph15):f2
(p10*7.2 ph16):f2
(p10*4.9 ph15):f2
(p10*7.4 ph16):f2
(p10*6.8 ph15):f2
(p10*7.0 ph16):f2
(p10*5.2 ph15):f2
(p10*5.4 ph16):f2
(p10*0.6 ph15):f2
(p10*4.5 ph16):f2
(p10*7.3 ph15):f2
(p10*5.1 ph16):f2
(p10*7.9 ph15):f2
lo to 8 times 3
(p5 ph12):f2     ;sl pulse
;total mixtime0.55*p6*391.2*loop
;*****  end DIPSI  ******
2u
1m dhi
(d21 p1*2 ph0) (p3*2 ph10):f2
2u
1m
(d21 p1 ph9) (p3 ph10):f2     ;inept
1.5m
(d21 p1*2 ph0) (p3*2 ph10):f2
d7
p7:e
3u adc
(p8 ph10):f2
(p9 ph11):f2
2u
2u dlo
        2u dl1
        d8 cpds     ;garp 60us
        do
rcyc=2 ph31
50m wr #0 if #0 zd
5m ip1
lo to 3 times 2
5m id0
5m ip31
5m ip31
lo to 4 times 128
5m rd0
5m ip5
lo to 5 times 2
5m id10
5m ip31
5m ip31
lo to 6 times 32
do
exit

ph0=0
ph1=0
ph2=0 0 0 0 0 0 0 0
   2 2 2 2 2 2 2 2
ph3=0 0 0 0 1 1 1 1
   2 2 2 2 3 3 3 3
ph4=1 3
ph5=0
ph6=0
ph7=0 0 1 1 2 2 3 3
ph9=0 0 0 0 0 0 0 0
   2 2 2 2 2 2 2 2
ph10=0
ph11=0 2
ph12=(360)5 5 5 5 185 185 185
185;should be adjusted
ph13=1 1 1 1 1 1 1 1
   3 3 3 3 3 3 3 3
ph15=(360)95          ;should be adjusted
ph16=(360)275         ;should be adjusted
ph21=(6)5
ph22=(6)4
ph23=(6)3
ph24=(6)2
ph25=(6)1
ph26=(6)0
ph31=0 2 2 0 2 0 0 2
      2 0 0 2 0 2 2 0
```

```
;hcchnoe4d.al
;4d noesy  (1H-13C-13C-1H)
;11/28/93 A. Lee,hcchnoe4d.abx as templat
;ref: Clore, Kay, Bax, Gronenborn
;Biochemistry, 1991, 30, pp. 12-18
;set d10 and d20 properly for 1/2 dw
p13=1m
p24=3.4m
d11=40m
d12=6m
d6=p3*4
d16=d6+10u
d8=aq
1 ze
 thi
2 d11
 d12*2
3 d12*6
4 d12*2
5 d12*3
6 d12*2
7 d12*3
8 d12 dlo hl2 o1
 p18 ph29 dl0 o2
 (p13 ph0 2u):f2
 d12 dhi hl1 o1
 p1 ph1
 3.4m
 d16
 (p3 ph11):f2
 d21
 p1*2 ph2
 d21
 (p3 ph0):f2                ;end of first hmqc
 3.4m o2
 d0
 (p3 ph22 2u p3*2 ph12 2u p3 ph22):f2
 d0
 (p2 ph3)                   ;trim
 2u
 p1 ph4
 40m o2
 (p3 ph0):f2
 60m
 10u dlo
 (p13 ph0):f2
 (p1 ph5)
 3.4m dhi
 10u
 (p3 ph13):f2               ;begin final hmqc
 d20
 (p1*2 ph6)
 d20
 (p3 ph0):f2
 p24:e o2
 d7                         ;pre-scan delay
 3u adc
 (p3 ph0):f2
 2u
 (p3 ph14):f2
 2u
 5u dlo
 5u dl1
 d8 cpds
 10u do
 rcyc=2 ph31
 d11 wr #0 if #0 zd
 d12 dp3
 d12 dp4        ;states part of states-tppi (t2)
 lo to 3 times 2
 d12 id0
 d12 rp3
 d12 rp4
 d12 ip1
 d12 ip1
 d12 ip31
 d12 ip31       ;tppi part of states-tppi (t2)
 lo to 4 times 64       ;64 complex in f1
 d12 rd0                ;reset d0
 d12 ip11               ;quad in t1 (13C)
lo to 5 times 2
 d12 ip31
 d12 id21
 d12 ip31
lo to 6 times 8
 d12 rd21
 d12 ip13               ;quad in t3 (13C)
lo to 7 times 2
 d12 ip31
 d12 id20
 d12 ip31
lo to 8 times 8
 exit
 ph0=0
 ph1=0
 ph2=0 0 0 0 1 1 1 1
 ph3=3
 ph4=0
 ph5=0
 ph6=0 0 0 0 1 1 1 1
 ph11=0 2
 ph12=1 1 3 3
 ph22=0 0 2 2
 ph13=0 0 2 2
 ph14=0 2
 ph29=0
 ph31=0 2 2 0
 ph7=2
```

Appendix B
Supplementary material to chapter 2

1. ^1H resonance assignments for ApaBPTI 10° C, pH 2.75.

Residue	NH	αH	βH,H'	other
Ty r 0		4.20	3.05, 3.16	ring (2,6) 7.11, (3,5) 6.85
Cys 1	8.68	4.75	2.53,2.83	
Asn 2	8.86	4.91	2.72, 2.82	NH_2 7.84, 7.22
Cys 3	9.07	4.70	3.06, 2.96	
Lys 4	8.08	4.29	1.70, 1.86	γCH_2 1.70, 1.86 δCH_2 1.42, 1.52 εCH_2 3.01, 3.01 NH_3 7.61
Ala 5	7.44	4.51	1.25	
Pro 6		4.72	2.07, 2.44	γCH_2 1.75, 1.95 δCH_2 3.50, 3.61
Glu 7	8.61	4.09	2.02, 2.09	γCH_2 2.50
Thr 8	7.88	4.56	4.54	γCH_3 1.19
Glu 9	9.21	4.10	2.52, 2.52	γCH_2 2.49
Asp 10	9.01	4.43	2.82, 2.82	
Cys 11	7.88	4.79	2.80, 3.12	
Met 12	8.54	3.95	2.35	γCH_2 2.81, 2.93 εCH_3 2.08
Arg 13	8.15	4.13	1.80, 1.99	γCH_2 1.62, 1.62 δCH_2 3.23 NH 7.28
Thr 14	8.31	4.35	4.07	γCH_3
Cys 15	8.31	4.55	2.97, 3.15	
Gly 16	7.87	4.06, 4.06		
Gly 17	8.18	4.02, 4.02		
Ala 18	8.23	4.38	1.43	

2. Distance restraints used in DGII calculations of ApaBPTI

```
!BIOSYM restraint 1
!
#remote_prochiral_center
!
#distance
1:CYS_2:CB      1:CYS_12:CB     3.72    3.99    1       1       1000
1:CYS_4:CB      1:CYS_16:CB     3.72    3.99    1       1       1000
1:THR_9:O       1:MET_13:HN     1.8     2.3     1       1       1000
1:THR_9:O       1:MET_13:N      2.8     3.3     1       1       1000
1:GLU_10:O      1:ARG+_14:HN    1.8     2.1     1       1       1000
1:GLU_10:O      1:ARG+_14:N     2.8     3.3     1       1       1000
1:ASP_11:O      1:THR_15:HN     1.8     2.1     1       1       1000
1:ASP_11:O      1:THR_15:N      2.8     3.3     1       1       1000
1:CYS_12:O      1:CYS_16:HN     1.8     2.1     1       1       1000
1:CYS_12:O      1:CYS_16:N      2.8     3.3     1       1       1000
!
#NOE_distance
1:TYRN_1:HA     1:TYRN_1:HB2    1.8     2.5     2.5     1       1       1000    0
1:TYRN_1:HA     1:TYRN_1:HB1    1.8     2.5     2.5     1       1       1000    0
1:TYRN_1:HD*    1:CYS_2:HN      1.8     4       4       1       1       1000    0
1:TYRN_1:HD*    1:TYRN_1:HA     1.8     6       4       1       1       1000    0
1:TYRN_1:HA     1:CYS_2:HN      1.8     2.5     2.5     1       1       1000    0
1:TYRN_1:HB1    1:CYS_2:HN      1.8     3.5     3.5     1       1       1000    0
1:TYRN_1:HB2    1:CYS_2:HN      1.8     3.5     3.5     1       1       1000    0
1:CYS_2:HA      1:ASN_3:HN      1.8     2.5     2.5     1       1       1000    0
1:CYS_2:HN      1:CYS_2:HB1     1.8     2.5     2.5     1       1       1000    0
1:CYS_2:HN      1:CYS_2:HB2     1.8     2.5     2.5     1       1       1000    0
1:CYS_2:HB2     1:ASN_3:HN      0.033   6.267   6.267   1       1       1000    0
1:CYS_2:HB2     1:CYS_12:HB1    -1      8.533   8.533   1       1       1000    0
1:ASN_3:HN      1:ASN_3:HB1     1.8     2.5     2.5     1       1       1000    0
1:ASN_3:HN      1:ASN_3:HB2     1.8     2.5     2.5     1       1       1000    0
1:ASN_3:HD22    1:ASN_3:HB2     1.8     2.5     2.5     1       1       1000    0
1:ASN_3:HD22    1:ASN_3:HB1     1.8     2.5     2.5     1       1       1000    0
1:ASN_3:HD21    1:ASN_3:HB2     1.8     4.5     4.5     1       1       1000    0
1:ASN_3:HD21    1:ASN_3:HB1     1.8     4.5     4.5     1       1       1000    0
1:ASN_3:HD21    1:ALA_6:HB*     1.8     5.5     4.5     1       1       1000    0
1:ASN_3:HD22    1:ALA_6:HB*     1.8     5.5     4.5     1       1       1000    0
1:ASN_3:HD21    1:LYS+_5:HB*    1.8     5.5     4.5     1       1       1000    0
1:ASN_3:HA      1:CYS_4:HN      1.8     2.5     2.5     1       1       1000    0
1:CYS_4:HN      1:CYS_4:HB2     1.8     2.5     2.5     1       1       1000    0
1:CYS_4:HN      1:CYS_4:HB1     1.8     2.5     2.5     1       1       1000    0
1:CYS_4:HN      1:LYS+_5:HN     1.8     2.5     2.5     1       1       1000    0
1:CYS_4:HB2     1:LYS+_5:HN     0.032   6.268   6.268   1       1       1000    0
1:CYS_4:HB1     1:CYS_16:HB1    1.8     6.03    6.03    1       1       1000    0
1:CYS_4:HA      1:CYS_16:HB1    0.038   6.262   6.262   1       1       1000    0
1:ASN_3:HB1     1:ALA_6:HN      0.057   6.243   6.243   1       1       1000    0
1:LYS+_5:HN     1:ALA_6:HB*     1.8     5.5     4.5     1       1       1000    0
1:LYS+_5:HN     1:LYS+_5:HA     1.8     2.5     2.5     1       1       1000    0
1:LYS+_5:HA     1:ALA_6:HN      1.8     3.5     3.5     1       1       1000    0
1:LYS+_5:HN     1:ALA_6:HN      1.8     2.5     2.5     1       1       1000    0
1:LYS+_5:HB*    1:LYS+_5:HN     1.8     3.5     2.5     1       1       1000    0
1:LYS+_5:HN     1:LYS+_5:HG1    1.8     4.5     4.5     1       1       1000    0
```

1:LYS+_5:HN	1:LYS+_5:HG2	1.8	4.5	4.5	1	1	1000	0
1:LYS+_5:HB*	1:ALA_6:HN	1.8	3.5	3.5	1	1	1000	0
1:ALA_6:HN	1:ALA_6:HA	1.8	2.5	2.5	1	1	1000	0
1:ALA_6:HN	1:PRO_7:HD2	1.8	3.5	3.5	1	1	1000	0
1:ALA_6:HN	1:PRO_7:HD1	1.8	3.5	3.5	1	1	1000	0
1:ALA_6:HN	1:ALA_6:HB*	1.8	3.5	2.5	1	1	1000	0
1:ALA_6:HA	1:PRO_7:HD2	1.8	2.5	2.5	1	1	1000	0
1:ALA_6:HA	1:PRO_7:HD1	1.8	2.5	2.5	1	1	1000	0
1:ALA_6:HB*	1:PRO_7:HD2	1.8	4.5	3.5	1	1	1000	0
1:ALA_6:HB*	1:PRO_7:HD1	1.8	4.5	3.5	1	1	1000	0
1:ALA_6:HB*	1:ASN_3:HB1	0.057	7.243	6.243	1	1	1000	0
1:PRO_7:HB*	1:MET_13:HA	1.8	4.5	3.5	1	1	1000	0
1:PRO_7:HB*	1:MET_13:HN	1.8	5.5	4.5	1	1	1000	0
1:PRO_7:HB*	1:GLU_8:HN	1.8	5.5	4.5	1	1	1000	0
1:GLU_8:HN	1:THR_9:HN	1.8	4.5	4.5	1	1	1000	0
1:GLU_8:HN	1:GLU_8:HG2	0.054	5.246	5.246	1	1	1000	0
1:GLU_8:HN	1:GLU_8:CB	1.8	3.5	3.5	1	1	1000	0
1:GLU_8:HN	1:GLU_8:HB2	1.8	3.5	3.5	1	1	1000	0
1:GLU_8:HN	1:GLU_8:HB1	1.8	3.5	3.5	1	1	1000	0
1:GLU_8:HB2	1:THR_9:HN	0.052	6.248	6.248	1	1	1000	0
1:GLU_8:HA	1:THR_9:HN	1.8	3.5	3.5	1	1	1000	0
1:THR_9:HA	1:THR_9:HG2*	1.8	3.5	2.5	1	1	1000	0
1:THR_9:HN	1:THR_9:HG2*	1.8	4.5	3.5	1	1	1000	0
1:THR_9:HA	1:GLU_10:HN	1.8	3.5	3.5	1	1	1000	0
1:THR_9:HB	1:GLU_10:HN	1.8	2.5	2.5	1	1	1000	0
1:THR_9:HA	1:CYS_12:HN	1.8	4.5	4.5	1	1	1000	0
1:THR_9:HG2*	1:GLU_10:HN	1.8	4.5	3.5	1	1	1000	0
1:THR_9:HG2*	1:ASP_11:HN	1.8	5.5	4.5	1	1	1000	0
1:GLU_10:HN	1:GLU_10:HA	1.8	2.5	2.5	1	1	1000	0
1:GLU_10:HN	1:GLU_10:HB2	1.8	2.5	2.5	1	1	1000	0
1:GLU_10:HN	1:GLU_10:HB1	1.8	2.5	2.5	1	1	1000	0
1:GLU_10:HN	1:GLU_10:HG*	1.8	4.5	3.5	1	1	1000	0
1:GLU_10:HA	1:ASP_11:HN	1.8	3.5	3.5	1	1	1000	0
1:GLU_10:HB1	1:ASP_11:HN	1.8	2.5	2.5	1	1	1000	0
1:GLU_10:HB2	1:ASP_11:HN	1.8	2.5	2.5	1	1	1000	0
1:GLU_10:HG*	1:ASP_11:HN	1.8	5.5	4.5	1	1	1000	0
1:GLU_10:HA	1:MET_13:HN	1.8	3.5	3.5	1	1	1000	0
1:GLU_10:HA	1:ARG+_14:HN	1.8	4.5	4.5	1	1	1000	0
1:ASP_11:HA	1:ASP_11:HN	1.8	2.5	2.5	1	1	1000	0
1:ASP_11:HN	1:ASP_11:HB*	1.8	3.5	2.5	1	1	1000	0
1:ASP_11:HA	1:CYS_12:HN	1.8	3.5	3.5	1	1	1000	0
1:ASP_11:HN	1:THR_9:HA	1.8	3.5	3.5	1	1	1000	0
1:GLU_10:HN	1:ASP_11:HN	1.8	2.5	2.5	1	1	1000	0
1:ASP_11:HN	1:CYS_12:HN	1.8	2.5	2.5	1	1	1000	0
1:ASP_11:HA	1:ARG+_14:HN	1.8	4.5	4.5	1	1	1000	0
1:ASP_11:HA	1:THR_15:HN	1.8	4.5	4.5	1	1	1000	0
1:ASP_11:HA	1:ARG+_14:HB*	1.8	4.5	3.5	1	1	1000	0
1:CYS_12:HB2	1:MET_13:HN	1.8	4.5	4.5	1	1	1000	0
1:CYS_12:HB1	1:MET_13:HN	1.8	4.5	4.5	1	1	1000	0
1:CYS_12:HN	1:CYS_12:HB1	1.8	2.5	2.5	1	1	1000	0
1:CYS_12:HN	1:CYS_12:HB2	1.8	2.5	2.5	1	1	1000	0
1:CYS_12:HN	1:MET_13:HN	1.8	2.5	2.5	1	1	1000	0
1:CYS_12:HA	1:THR_15:HN	1.8	4.5	4.5	1	1	1000	0
1:MET_13:HN	1:MET_13:HA	1.8	2.5	2.5	1	1	1000	0

1:MET_13:HA	1:MET_13:HB*	1.8	3.5	2.5	1	1	1000	0
1:MET_13:HA	1:MET_13:HG*	1.8	3.5	2.5	1	1	1000	0
1:MET_13:HA	1:ARG+_14:HN	1.8	3.5	3.5	1	1	1000	0
1:MET_13:HB*	1:ARG+_14:HN	1.8	3.5	2.5	1	1	1000	0
1:MET_13:HG*	1:ARG+_14:HN	1.8	5.5	4.5	1	1	1000	0
1:MET_13:HA	1:CYS_16:HN	1.8	3.5	3.5	1	1	1000	0
1:MET_13:HN	1:ARG+_14:HN	1.8	2.5	2.5	1	1	1000	0
1:MET_13:HA	1:CYS_16:HB1	1.8	4.5	4.5	1	1	1000	0
1:MET_13:HA	1:CYS_16:HB2	1.8	4.5	4.5	1	1	1000	0
1:ARG+_14:HN	1:ARG+_14:HA	1.8	2.5	2.5	1	1	1000	0
1:ARG+_14:HN	1:ARG+_14:HB*	1.8	3.5	2.5	1	1	1000	0
1:ARG+_14:HN	1:ARG+_14:HG*	1.8	4.5	3.5	1	1	1000	0
1:ARG+_14:HN	1:ARG+_14:HD*	1.8	5.5	4.5	1	1	1000	0
1:ARG+_14:HE	1:ARG+_14:HB*	1.8	4.5	3.5	1	1	1000	0
1:ARG+_14:HA	1:THR_15:HN	1.8	3.5	3.5	1	1	1000	0
1:ARG+_14:HN	1:THR_15:HN	1.8	2.5	2.5	1	1	1000	0
1:ARG+_14:HB*	1:THR_15:HN	1.8	3.5	2.5	1	1	1000	0
1:ARG+_14:HA	1:GLY_17:HN	1.8	3.5	3.5	1	1	1000	0
1:ARG+_14:HA	1:GLY_18:HN	1.8	3.5	3.5	1	1	1000	0
1:THR_15:HN	1:THR_15:HA	1.8	2.5	2.5	1	1	1000	0
1:THR_15:HN	1:THR_15:HB	1.8	3.5	3.5	1	1	1000	0
1:THR_15:HA	1:CYS_16:HN	1.8	3.5	3.5	1	1	1000	0
1:THR_15:HB	1:CYS_16:HN	1.8	3.5	3.5	1	1	1000	0
1:THR_15:HN	1:THR_15:HG2*	1.8	4.5	3.5	1	1	1000	0
1:THR_15:HG2*	1:CYS_16:HN	1.8	5.5	4.5	1	1	1000	0
1:CYS_16:HN	1:CYS_16:HA	1.8	2.5	2.5	1	1	1000	0
1:CYS_16:HN	1:CYS_16:HB1	1.8	3.5	3.5	1	1	1000	0
1:CYS_16:HN	1:CYS_16:HB2	1.8	3.5	3.5	1	1	1000	0
1:CYS_16:HA	1:GLY_17:HN	1.8	3.5	3.5	1	1	1000	0
1:CYS_16:HB1	1:GLY_17:HN	1.8	3.5	3.5	1	1	1000	0
1:CYS_16:HB2	1:GLY_17:HN	1.8	3.5	3.5	1	1	1000	0
1:THR_15:HG2*	1:CYS_16:HA	1.8	4.5	4.5	1	1	1000	0
1:GLY_17:HA*	1:GLY_17:HN	1.8	3.5	2.5	1	1	1000	0
1:CYS_16:HN	1:GLY_17:HN	1.8	2.5	2.5	1	1	1000	0
1:GLY_17:HN	1:GLY_18:HN	1.8	2.5	2.5	1	1	1000	0
1:GLY_18:HN	1:GLY_18:HA*	1.8	3.5	2.5	1	1	1000	0

!
#mixing_times
-9.99E+02

3. RMSDs of 7 best (of 10) ApaBPTI DGII structures

# in ab.arc	vs structure 1 rmsd bb:2-17	ab:10-17 vs 5pti:49-56 rmsd bb:	rmsd all heavy
1	0.000	0.973	2.147
3	0.549	1.015	1.905
4	0.605	1.133	1.906
5	0.189	0.948	1.914
7	0.171	0.952	2.145
8	0.303	1.008	2.264
9	0.133	0.942	1.778
average	0.28 ± 0.22	1.00 ± 0.07	2.01 ± 0.18

Appendix C
Supplementary material to chapter 3

Fur (88-147) ^{15}N & ^{1}H NMR Assignments

Residue	Strip	^{15}N	NH	αH	βH	
Leu 90	7	127.69	9.47	5.31		
Ile 91	12	127.69	9.07	5.13		
Cys 92	37	126.96	8.29		3.11	
Leu 93	9	128.5	9.34		1.70, 1.45	
Asp 94	5	121.94	9.60		2.78, 2.75	
Cys 95						
Gly 96	47	114.73	8.09			
Lys 97	20	123.80	8.59			
Val 98	48	122.89	8.09	4.09		1.46
Ile 99	1	131.43	9.89		1.60	
Glu 100	19	127.47	7.93	4.95	1.90	
Ala 108	44	120.76	8.12	4.04	1.48	
Arg 109	60	118.31	7.78	4.16		
Gln 110	33	120.21	8.323	4.94		
Arg 111	46	124.28	8.08			
Glu 112	56	120.26	7.88	4.11		
Ile 113	35					
Ala 114	58	120.26	7.88			
Ala 115	40	121.00	8.26			
Lys 116	50	120.83	8.02	4.18	1.89	0.91
His 117	69	114.03	7.33		2.61, 2.58	
Gly 118	45	109.06	8.12	4.15		
Ile 119	31	122.17	8.36	4.16	1.91	
Arg 120	6	129.22	9.49			
Leu 121	24	128.36	8.50			
Thr 122	2	116.89	9.76	4.20		
Asn 123	36	118.95	8.29	5.09		
His 124	18	114.75	8.69			
Ser 125	8	115.25	9.40			
Leu 126	11	129.96	9.18			
Tyr 127	13	126.01	9.04	5.70		
Leu 128	10	121.92	9.19			
Gly 130	43	106.96	8.14			
Cys 132	3	123.96	9.63	4.52		
Ala 133	17	130.85	8.71	4.14	1.32	
Glu 134	53	116.55	7.99	4.52	1.91	
Gly 135	22	111.57	8.58	3.98, 3.70		
Asp 136	25	119.99	8.44	4.96	2.84	
Cys 137	54	127.47	7.93	4.45	2.92	
Arg 138	39	122.56	8.23	4.09		
Glu 139	62	115.32	7.65	4.44	1.85, 2.20	
Asp 140	68	118.93	7.35	4.44	2.73, 2.51	
Glu 141	14	125.38	9.00		1.87	
His 142	23	117.74	8.57			
Ala 143	67	125.16	7.42	3.91	1.19	
His 144	15	111.72	8.88		3.90	
Glu 145	65	119.2	7.46	4.39	2.10, 2.43	
Gly 146	16	109.89	8.79	3.77, 3.55		
Lys 147	54	127.47	7.93	4.20	1.67, 1.83, 1.38	

Appendix D
Supplementary material to chapter 4

Restraint files for DGSA calculations of NTRC receiver domain structure.

X-PLOR input file dihe_tbl; contains all ϕ angle restraints.

```
restraints dihedral reset
assign (resid 7 and name HN) (resid 7 and name N)
       (resid 7 and name CA) (resid 7 and name HA) 1.0 -120.0 40.0  2
assign (resid 8 and name HN) (resid 8 and name N)
       (resid 8 and name CA) (resid 8 and name HA) 1.0 -120.0 40.0  2
assign (resid 18 and name HN) (resid 18 and name N)
       (resid 18 and name CA) (resid 18 and name HA) 1.0 -60.0 30.0  2
assign (resid 20 and name HN) (resid 20 and name N)
       (resid 20 and name CA) (resid 20 and name HA) 1.0 -60.0 30.0  2
assign (resid 29 and name HN) (resid 29 and name N)
       (resid 29 and name CA) (resid 29 and name HA) 1.0 -120.0 40.0  2
assign (resid 30 and name HN) (resid 30 and name N)
       (resid 30 and name CA) (resid 30 and name HA) 1.0 -120.0 40.0  2
assign (resid 31 and name HN) (resid 31 and name N)
       (resid 31 and name CA) (resid 31 and name HA) 1.0 -120.0 40.0  2
assign (resid 38 and name HN) (resid 38 and name N)
       (resid 38 and name CA) (resid 38 and name HA) 1.0 -60.0 30.0  2
assign (resid 45 and name HN) (resid 45 and name N)
       (resid 45 and name CA) (resid 45 and name HA) 1.0 -120.0 40.0  2
assign (resid 47 and name HN) (resid 47 and name N)
       (resid 47 and name CA) (resid 47 and name HA) 1.0 -120.0 40.0  2
assign (resid 51 and name HN) (resid 51 and name N)
       (resid 51 and name CA) (resid 51 and name HA) 1.0 -120.0 40.0  2
assign (resid 69 and name HN) (resid 69 and name N)
       (resid 69 and name CA) (resid 69 and name HA) 1.0 -60.0 30.0  2
assign (resid 71 and name HN) (resid 71 and name N)
       (resid 71 and name CA) (resid 71 and name HA) 1.0 -60.0 30.0  2
assign (resid 87 and name HN) (resid 87 and name N)
       (resid 87 and name CA) (resid 87 and name HA) 1.0 -60.0 30.0  2
assign (resid 91 and name HN) (resid 91 and name N)
       (resid 91 and name CA) (resid 91 and name HA) 1.0 -60.0 30.0  2
assign (resid 112 and name HN) (resid 112 and name N)
       (resid 112 and name CA) (resid 112 and name HA) 1.0 -60.0 30.0  2
assign (resid 114 and name HN) (resid 114 and name N)
       (resid 114 and name CA) (resid 114 and name HA) 1.0 -60.0 30.0  2
assign (resid 116 and name HN) (resid 116 and name N)
       (resid 116 and name CA) (resid 116 and name HA) 1.0 -60.0 30.0  2
assign (resid 119 and name HN) (resid 119 and name N)
       (resid 119 and name CA) (resid 119 and name HA) 1.0 -60.0 30.0  2
end
```

X-PLOR file noe_tbl; contains distance restraints for NOEs and H-Bonds.

```
!hydrogen bond restraints
```

assign (resid 6 and name O) (resid 31 and name HN) 2.0 0.2 0.3
assign (resid 6 and name O) (resid 31 and name N) 3.0 0.2 0.3
assign (resid 4 and name O) (resid 29 and name HN) 2.0 0.2 0.3{rnd11added}
assign (resid 4 and name O) (resid 29 and name N) 3.0 0.2 0.3{rnd11added}
assign (resid 49 and name OD2) (resid 5 and name HN) 2.0 0.2 0.3
assign (resid 49 and name OD2) (resid 5 and name N) 3.0 0.2 0.3
assign (resid 29 and name O) (resid 6 and name HN) 2.0 0.2 0.3
assign (resid 29 and name O) (resid 6 and name N) 3.0 0.2 0.3
assign (resid 31 and name O) (resid 8 and name HN) 2.0 0.2 0.3
assign (resid 31 and name O) (resid 8 and name N) 3.0 0.2 0.3
assign (resid 8 and name O) (resid 33 and name HN) 2.0 0.2 0.3
assign (resid 8 and name O) (resid 33 and name N) 3.0 0.2 0.3
assign (resid 9 and name O) (resid 54 and name HN) 2.0 0.2 0.3}
assign (resid 9 and name O) (resid 54 and name N) 3.0 0.2 0.3
assign (resid 5 and name O) (resid 50 and name HN) 2.0 0.2 0.3
assign (resid 5 and name O) (resid 50 and name N) 3.0 0.2 0.3
assign (resid 5 and name O) (resid 49 and name HN) 2.0 0.2 0.3
assign (resid 5 and name O) (resid 49 and name N) 3.0 0.2 0.3
assign (resid 50 and name O) (resid 7 and name HN) 2.0 0.2 0.3
assign (resid 50 and name O) (resid 7 and name N) 3.0 0.2 0.3
assign (resid 102 and name O) (resid 82 and name HN) 2.0 0.2 0.3
assign (resid 102 and name O) (resid 82 and name N) 3.0 0.2 0.3
assign (resid 7 and name O) (resid 52 and name HN) 2.0 0.2 0.3
assign (resid 7 and name O) (resid 52 and name N) 3.0 0.2 0.3
assign (resid 52 and name O) (resid 9 and name HN) 2.0 0.2 0.3
assign (resid 52 and name O) (resid 9 and name N) 3.0 0.2 0.3
assign (resid 33 and name O) (resid 10 and name HN) 2.0 0.2 0.3
assign (resid 33 and name O) (resid 10 and name N) 3.0 0.2 0.3
assign (resid 35 and name O) (resid 39 and name N) 3.0 0.2 0.3
assign (resid 35 and name O) (resid 39 and name HN) 2.0 0.2 0.3
assign (resid 38 and name O) (resid 42 and name N) 3.0 0.2 0.3
assign (resid 38 and name O) (resid 42 and name HN) 2.0 0.2 0.3
assign (resid 39 and name O) (resid 43 and name N) 3.0 0.2 0.3
assign (resid 39 and name O) (resid 43 and name HN) 2.0 0.2 0.3
assign (resid 98 and name O) (resid 78 and name HN) 2.0 0.2 0.3
assign (resid 98 and name O) (resid 78 and name N) 3.0 0.2 0.3
assign (resid 77 and name O) (resid 51 and name HN) 2.0 0.2 0.3
assign (resid 77 and name O) (resid 51 and name N) 3.0 0.2 0.3
assign (resid 51 and name O) (resid 79 and name HN) 2.0 0.2 0.3
assign (resid 51 and name O) (resid 79 and name N) 3.0 0.2 0.3
assign (resid 79 and name O) (resid 53 and name HN) 2.0 0.2 0.3
assign (resid 79 and name O) (resid 53 and name N) 3.0 0.2 0.3
assign (resid 53 and name O) (resid 81 and name HN) 2.0 0.2 0.3
assign (resid 53 and name O) (resid 81 and name N) 3.0 0.2 0.3
assign (resid 15 and name O) (resid 19 and name HN) 2.0 0.2 0.3
assign (resid 15 and name O) (resid 19 and name N) 3.0 0.2 0.3
assign (resid 16 and name O) (resid 20 and name HN) 2.0 0.2 0.3
assign (resid 16 and name O) (resid 20 and name N) 3.0 0.2 0.3
assign (resid 18 and name O) (resid 22 and name HN) 2.0 0.2 0.3
assign (resid 18 and name O) (resid 22 and name N) 3.0 0.2 0.3
assign (resid 19 and name O) (resid 23 and name HN) 2.0 0.2 0.3
assign (resid 19 and name O) (resid 23 and name N) 3.0 0.2 0.3
assign (resid 20 and name O) (resid 24 and name HN) 2.0 0.2 0.3
assign (resid 20 and name O) (resid 24 and name N) 3.0 0.2 0.3

assign (resid 23 and name O) (resid 28 and name HN) 2.0 0.2 0.3
assign (resid 23 and name O) (resid 28 and name N) 3.0 0.2 0.3
assign (resid 65 and name O) (resid 69 and name HN) 2.0 0.2 0.3
assign (resid 65 and name O) (resid 69 and name N) 3.0 0.2 0.3
assign (resid 80 and name O) (resid 102 and name HN) 2.0 0.2 0.3
assign (resid 80 and name O) (resid 102 and name N) 3.0 0.2 0.3
assign (resid 67 and name O) (resid 71 and name HN) 2.0 0.2 0.3
assign (resid 67 and name O) (resid 71 and name N) 3.0 0.2 0.3
assign (resid 68 and name O) (resid 72 and name HN) 2.0 0.2 0.3
assign (resid 68 and name O) (resid 72 and name N) 3.0 0.2 0.3
assign (resid 86 and name O) (resid 90 and name HN) 2.0 0.2 0.3
assign (resid 86 and name O) (resid 90 and name N) 3.0 0.2 0.3
assign (resid 108 and name O) (resid 112 and name HN) 2.0 0.2 0.3
assign (resid 108 and name O) (resid 112 and name N) 3.0 0.2 0.3
assign (resid 109 and name O) (resid 113 and name HN) 2.0 0.2 0.3
assign (resid 109 and name O) (resid 113 and name N) 3.0 0.2 0.3
assign (resid 110 and name O) (resid 114 and name HN) 2.0 0.2 0.3
assign (resid 110 and name O) (resid 114 and name N) 3.0 0.2 0.3
assign (resid 111 and name O) (resid 115 and name HN) 2.0 0.2 0.3
assign (resid 111 and name O) (resid 115 and name N) 3.0 0.2 0.3
assign (resid 112 and name O) (resid 116 and name HN) 2.0 0.2 0.3
assign (resid 112 and name O) (resid 116 and name N) 3.0 0.2 0.3
assign (resid 113 and name O) (resid 117 and name HN) 2.0 0.2 0.3
assign (resid 113 and name O) (resid 117 and name N) 3.0 0.2 0.3
assign (resid 115 and name O) (resid 119 and name HN) 2.0 0.2 0.3
assign (resid 115 and name O) (resid 119 and name N) 3.0 0.2 0.3
assign (resid 19 and name O) (resid 23 and name HN) 2.0 0.2 0.3
assign (resid 19 and name O) (resid 23 and name N) 3.0 0.2 0.3
!
!NOE-based restraints
assign (resid 2 and name HG*) (resid 2 and name HA) 2.0 0.2 2.5
assign (resid 2 and name HA) (resid 3 and name HN) 2.0 0.2 1.5
assign (resid 2 and name HB*) (resid 3 and name HN) 2.0 0.2 2.413
assign (resid 2 and name HB*) (resid 28 and name HD*) 2.0 0.2 6.9
assign (resid 2 and name HG*) (resid 28 and name HD*) 2.0 0.2 6.9
assign (resid 3 and name HG*) (resid 3 and name HA) 2.0 0.2 1.7
assign (resid 3 and name HA) (resid 4 and name HN) 2.0 0.2 0.5
assign (resid 3 and name HB*) (resid 4 and name HN) 2.0 0.2 2.5
assign (resid 3 and name HN) (resid 4 and name HN) 2.0 0.2 2.5
assign (resid 3 and name HG*) (resid 28 and name HD*) 2.0 0.2 5
assign (resid 3 and name HA) (resid 116 and name HB*) 2.0 0.2 5.0
assign (resid 4 and name HA*) (resid 5 and name HN) 2 0.2 1.5
assign (resid 4 and name HN) (resid 5 and name HN) 2.0 0.2 3.0
assign (resid 4 and name HA*) (resid 28 and name HD*) 2.0 0.2 4.6
assign (resid 4 and name HA*) (resid 28 and name HD2*) 2.0 0.2 4.6
assign (resid 4 and name HN) (resid 28 and name HA) 2.0 0.2 1.5
assign (resid 4 and name HN) (resid 28 and name HD*) 2.0 0.2 4.4
assign (resid 4 and name HN) (resid 29 and name HN) 2.0 0.2 3.0
assign (resid 4 and name HA1) (resid 119 and name HD1*) 2.0 0.2 4.5
assign (resid 4 and name HA2) (resid 119 and name HD1*) 2.0 0.2 4.5
assign (resid 5 and name HA) (resid 5 and name HD*) 2.0 0.2 3.0
assign (resid 5 and name HA) (resid 5 and name HD*) 2.0 0.2 4.5
assign (resid 5 and name HB) (resid 5 and name HD*) 2.0 0.2 2.2
assign (resid 5 and name HG11) (resid 5 and name HA) 2.0 0.2 3.0

assign (resid 5 and name HG12) (resid 5 and name HA) 2.0 0.2 3.0
assign (resid 5 and name HA) (resid 6 and name HN) 2.0 0.2 1.5
assign (resid 5 and name HA) (resid 29 and name HB) 2.0 0.2 0.5
assign (resid 5 and name HA) (resid 29 and name HG1*) 2.0 0.2 3.0
assign (resid 5 and name HA) (resid 29 and name HG2*) 2.0 0.2 3.0
assign (resid 5 and name HA) (resid 29 and name HN) 2.0 0.2 2.5
assign (resid 5 and name HD1*) (resid 29 and name HB) 2.0 0.2 2.8
assign (resid 5 and name HD1*) (resid 29 and name HG2*) 2.0 0.2 4.0
assign (resid 5 and name HG1*) (resid 29 and name HG2*) 2.0 0.2 5.5
assign (resid 5 and name HG11) (resid 29 and name HB) 2.0 0.2 3.0
assign (resid 5 and name HG12) (resid 29 and name HB) 2.0 0.2 3.0
assign (resid 5 and name HA) (resid 30 and name HA) 2.0 0.2 3
assign (resid 5 and name HG1*) (resid 48 and name HD*) 2.0 0.2 5.0
assign (resid 5 and name HB) (resid 49 and name HN) 2.0 0.2 1.5
assign (resid 5 and name HD*) (resid 49 and name HN) 2.0 0.2 3.0
assign (resid 5 and name HN) (resid 49 and name HB*) 2.0 0.2 2.5
assign (resid 5 and name HN) (resid 49 and name HN) 2.0 0.2 2.5
assign (resid 5 and name HB) (resid 50 and name HN) 2.0 0.2 2.5
assign (resid 6 and name HB) (resid 6 and name HN) 2.0 0.2 0.5
assign (resid 6 and name HG*) (resid 6 and name HA) 2.0 0.2 3.6
assign (resid 6 and name HA) (resid 7 and name HN) 2.0 0.2 0.5
assign (resid 6 and name HG*) (resid 7 and name HD1*) 2.0 0.2 5.9
assign (resid 6 and name HG*) (resid 8 and name HB) 2.0 0.2 3.6
assign (resid 6 and name HB) (resid 23 and name HD*) 2.0 0.2 4.2
assign (resid 6 and name HA) (resid 28 and name HD*) 2.0 0.2 5.9
assign (resid 6 and name HB) (resid 28 and name HD*) 2.0 0.2 4.4
assign (resid 6 and name HG*) (resid 28 and name HB*) 2.0 0.2 6.9
assign (resid 6 and name HN) (resid 29 and name HN) 2.0 0.2 3
assign (resid 6 and name HB) (resid 30 and name HA) 2.0 0.2 3
assign (resid 6 and name HG*) (resid 30 and name HA) 2.0 0.2 4.4
assign (resid 6 and name HG*) (resid 30 and name HB1) 2.0 0.2 6.9
assign (resid 6 and name HG*) (resid 30 and name HB2) 2.0 0.2 6.9
assign (resid 6 and name HN) (resid 30 and name HA) 2.0 0.2 1.5
assign (resid 6 and name HG*) (resid 31 and name HN) 2.0 0.2 5.9
assign (resid 6 and name HN) (resid 31 and name HN) 2.0 0.2 3
assign (resid 6 and name HA) (resid 50 and name HN) 2.0 0.2 2.5
assign (resid 6 and name HA) (resid 50 and name HB) 2.0 0.2 1.5
assign (resid 7 and name HA) (resid 7 and name HD1) 2.0 0.2 1.7
assign (resid 7 and name HA) (resid 8 and name HN) 2.0 0.2 0.5
assign (resid 7 and name HB*) (resid 8 and name HN) 2.0 0.2 3.5
assign (resid 7 and name HH2) (resid 29 and name HG2*) 2.0 0.2 4.5
assign (resid 7 and name HA) (resid 30 and name HA) 2.0 0.2 3
assign (resid 7 and name HA) (resid 31 and name HB) 2.0 0.2 3
assign (resid 7 and name HA) (resid 31 and name HG2*) 2.0 0.2 4.5
assign (resid 7 and name HA) (resid 31 and name HN) 2.0 0.2 1.5
assign (resid 7 and name HD1) (resid 31 and name HB) 2.0 0.2 1.5
assign (resid 7 and name HD1) (resid 31 and name HG2*) 2.0 0.2 3.0
assign (resid 7 and name HA) (resid 32 and name HA) 2.0 0.2 3
assign (resid 7 and name HE3) (resid 39 and name HG1*) 2.0 0.2 4.4
assign (resid 7 and name HE3) (resid 39 and name HG2*) 2.0 0.2 4.4
assign (resid 7 and name HE3) (resid 43 and name HD*) 2.0 0.2 5.9
assign (resid 7 and name HN) (resid 50 and name HN) 2.0 0.2 3
assign (resid 7 and name HA) (resid 51 and name HD*) 2.0 0.2 5.8
assign (resid 7 and name HB*) (resid 51 and name HD*) 2.0 0.2 5.4

assign (resid 7 and name HE3) (resid 51 and name HD1*) 2.0 0.2 4.5
assign (resid 7 and name HE3) (resid 51 and name HD2*) 2.0 0.2 4.5
assign (resid 7 and name HN) (resid 51 and name HA) 2.0 0.2 1.5
assign (resid 7 and name HN) (resid 51 and name HD*) 2.0 0.2 5.9
assign (resid 7 and name HN) (resid 51 and name HG) 2.0 0.2 3.0
assign (resid 7 and name HN) (resid 52 and name HN) 2.0 0.2 5.8
assign (resid 8 and name HB) (resid 8 and name HN) 2.0 0.2 0.5
assign (resid 8 and name HA) (resid 9 and name HN) 2.0 0.2 0.5
assign (resid 8 and name HB) (resid 32 and name HA) 2.0 0.2 1.5
assign (resid 8 and name HB) (resid 32 and name HB) 2.0 0.2 3
assign (resid 8 and name HB) (resid 32 and name HG2*) 2.0 0.2 2.0
assign (resid 8 and name HN) (resid 32 and name HA) 2.0 0.2 1.5
assign (resid 8 and name HN) (resid 33 and name HN) 2.0 0.2 3
assign (resid 8 and name HA) (resid 51 and name HA) 2.0 0.2 3
assign (resid 8 and name HA) (resid 52 and name HB1) 2.0 0.2 3
assign (resid 8 and name HA) (resid 52 and name HB2) 2.0 0.2 3
assign (resid 8 and name HA) (resid 52 and name HN) 2.0 0.2 1.5
assign (resid 8 and name HA) (resid 53 and name HA) 2.0 0.2 3
assign (resid 9 and name HA) (resid 9 and name HG*) 2.0 0.2 3.6
assign (resid 9 and name HB) (resid 9 and name HN) 2.0 0.2 1.5
assign (resid 9 and name HA) (resid 10 and name HN) 2.0 0.2 0.5
assign (resid 9 and name HA) (resid 32 and name HA) 2.0 0.2 3
assign (resid 9 and name HA) (resid 33 and name HN) 2.0 0.2 2.5
assign (resid 9 and name HG*) (resid 33 and name HB1) 2.0 0.2 5.9
assign (resid 9 and name HG*) (resid 33 and name HB2) 2.0 0.2 5.9
assign (resid 9 and name HG1*) (resid 35 and name HA) 2.0 0.2 4.5
assign (resid 9 and name HG2*) (resid 35 and name HA) 2.0 0.2 4.5
assign (resid 9 and name HB) (resid 39 and name HG*) 2.0 0.2 5.9
assign (resid 9 and name HG*) (resid 39 and name HG*) 2.0 0.2 6
assign (resid 9 and name HN) (resid 52 and name HN) 2.0 0.2 3
assign (resid 9 and name HB) (resid 53 and name HA) 2.0 0.2 3
assign (resid 9 and name HG*) (resid 53 and name HA) 2.0 0.2 4.4
assign (resid 9 and name HN) (resid 53 and name HA) 2.0 0.2 1.5
assign (resid 9 and name HG*) (resid 57 and name HE*) 2.0 0.2 5.4
assign (resid 10 and name HN) (resid 10 and name HB1) 2.0 0.2 1.5
assign (resid 10 and name HN) (resid 10 and name HB2) 2.0 0.2 1.5
assign (resid 10 and name HA) (resid 11 and name HN) 2.0 0.2 1.5
assign (resid 10 and name HB1) (resid 15 and name HB) 2.0 0.2 3.0
assign (resid 10 and name HB2) (resid 15 and name HB) 2.0 0.2 3.0
assign (resid 10 and name HN) (resid 33 and name HN) 2.0 0.2 3.0
assign (resid 10 and name HN) (resid 34 and name HA) 2.0 0.2 1.5
assign (resid 10 and name HB*) (resid 39 and name HG*) 2.0 0.2 6.9
assign (resid 11 and name HN) (resid 11 and name HB1) 2.0 0.2 1.5
assign (resid 11 and name HN) (resid 11 and name HB2) 2.0 0.2 1.5
assign (resid 11 and name HN) (resid 12 and name HN) 2.0 0.2 1.5
assign (resid 11 and name HB*) (resid 15 and name HG2*) 2.0 0.2 4.0
assign (resid 11 and name HA) (resid 56 and name HD*) 2.0 0.2 4
assign (resid 12 and name HA) (resid 13 and name HN) 2.0 0.2 0.5
assign (resid 12 and name HB*) (resid 15 and name HG1*) 2.0 0.2 5.5
assign (resid 12 and name HB*) (resid 15 and name HG2*) 2.0 0.2 5.5
assign (resid 12 and name HB*) (resid 15 and name HN) 2.0 0.2 4.0
assign (resid 13 and name HN) (resid 14 and name HN) 2.0 0.2 3.0
assign (resid 14 and name HB*) (resid 14 and name HN) 2.0 0.2 2.5
assign (resid 14 and name HN) (resid 15 and name HN) 2.0 0.2 0.7

assign (resid 14 and name HN) (resid 16 and name HN) 2.0 0.2 2.5
assign (resid 14 and name HA) (resid 17 and name HB1) 2.0 0.2 1.5
assign (resid 14 and name HA) (resid 17 and name HB2) 2.0 0.2 1.5
assign (resid 14 and name HA) (resid 17 and name HN) 2.0 0.2 1.5
assign (resid 14 and name HA) (resid 18 and name HN) 2.0 0.2 2.5
assign (resid 15 and name HB) (resid 15 and name HN) 2.0 0.2 0.5
assign (resid 15 and name HD1*) (resid 15 and name HA) 2.0 0.2 2.2
assign (resid 15 and name HB) (resid 16 and name HN) 2.0 0.2 1.5
assign (resid 15 and name HN) (resid 16 and name HN) 2.0 0.2 1.5
assign (resid 15 and name HA) (resid 17 and name HN) 2.0 0.2 3.0
assign (resid 15 and name HA) (resid 18 and name HB) 2.0 0.2 1.5
assign (resid 15 and name HA) (resid 18 and name HG*) 2.0 0.2 3.6
assign (resid 15 and name HA) (resid 18 and name HN) 2.0 0.2 1.5
assign (resid 15 and name HD1*) (resid 18 and name HG*) 2.0 0.2 5.4
assign (resid 15 and name HG2*) (resid 19 and name HD*) 2.0 0.2 4.6
assign (resid 15 and name HB) (resid 34 and name HA) 2.0 0.2 2.5
assign (resid 16 and name HA) (resid 17 and name HN) 2.0 0.2 5.0
assign (resid 16 and name HB*) (resid 17 and name HN) 2.0 0.2 2.5
assign (resid 16 and name HN) (resid 17 and name HN) 2.0 0.2 1.5
assign (resid 16 and name HN) (resid 18 and name HN) 2.0 0.2 2.5
assign (resid 16 and name HD*) (resid 20 and name HB*) 2.0 0.2 5
assign (resid 16 and name HA) (resid 32 and name HA) 2.0 0.2 3.0
assign (resid 16 and name HA) (resid 32 and name HB) 2.0 0.2 3.0
assign (resid 16 and name HD*) (resid 32 and name HA) 2.0 0.2 4.0
assign (resid 16 and name HD*) (resid 32 and name HB) 2.0 0.2 4
assign (resid 17 and name HA) (resid 17 and name HD1) 2.0 0.2 3.0
assign (resid 17 and name HB*) (resid 17 and name HN) 2.0 0.2 1.5
assign (resid 17 and name HA) (resid 18 and name HN) 2.0 0.2 3.0
assign (resid 17 and name HB*) (resid 18 and name HN) 2.0 0.2 2.5
assign (resid 17 and name HE3) (resid 18 and name HA) 2.0 0.2 3.0
assign (resid 17 and name HE3) (resid 18 and name HG*) 2.0 0.2 5.9
assign (resid 17 and name HN) (resid 18 and name HN) 2.0 0.2 1.5
assign (resid 17 and name HA) (resid 20 and name HB*) 2.0 0.2 4
assign (resid 17 and name HA) (resid 20 and name HG*) 2.0 0.2 4
assign (resid 17 and name HA) (resid 20 and name HN) 2.0 0.2 2.5
assign (resid 18 and name HB) (resid 18 and name HN) 2.0 0.2 0.5
assign (resid 18 and name HG*) (resid 18 and name HA) 2.0 0.2 3.6
assign (resid 18 and name HG*) (resid 18 and name HN) 2.0 0.2 5.9
assign (resid 18 and name HG*) (resid 18 and name HN) 2.0 0.2 5.9
assign (resid 18 and name HB) (resid 19 and name HN) 2.0 0.2 1.5
assign (resid 18 and name HA) (resid 21 and name HB*) 2.0 0.2 4.0
assign (resid 19 and name HA) (resid 19 and name HD*) 2.0 0.2 3.6
assign (resid 19 and name HA) (resid 20 and name HN) 2.0 0.2 1.5
assign (resid 19 and name HN) (resid 20 and name HN) 2.0 0.2 1.5
assign (resid 19 and name HA) (resid 22 and name HB*) 2.0 0.2 2.2
assign (resid 19 and name HD*) (resid 22 and name HB*) 2.0 0.2 5.4
assign (resid 19 and name HD*) (resid 22 and name HB*) 2.0 0.2 6.5
assign (resid 19 and name HA) (resid 108 and name HD1*) 2.0 0.2 3.0
assign (resid 19 and name HD*) (resid 108 and name HA) 2.0 0.2 4.5
assign (resid 19 and name HD*) (resid 108 and name HD1*) 2.0 0.2 4.5
assign (resid 19 and name HD*) (resid 108 and name HG1*) 2.0 0.2 6.9
assign (resid 20 and name HN) (resid 20 and name HB*) 2.0 0.2 1.389
assign (resid 20 and name HN) (resid 21 and name HN) 2.0 0.2 1.5
assign (resid 20 and name HA) (resid 23 and name HB*) 2.0 0.2 0.3

assign (resid 20 and name HA) (resid 23 and name HD*) 2.0 0.2 5.9
assign (resid 20 and name HG*) (resid 23 and name HD*) 2.0 0.2 6.9
assign (resid 20 and name HA) (resid 24 and name HN) 2.0 0.2 3.0
assign (resid 20 and name HA) (resid 30 and name HB1) 2.0 0.2 3.0
assign (resid 20 and name HA) (resid 30 and name HB2) 2.0 0.2 3.0
assign (resid 20 and name HB*) (resid 30 and name HB*) 2.0 0.2 5.0
assign (resid 20 and name HG*) (resid 30 and name HB1) 2.0 0.2 4
assign (resid 20 and name HG*) (resid 30 and name HB2) 2.0 0.2 4
assign (resid 20 and name HB*) (resid 32 and name HG1*) 2.0 0.2 3.2
assign (resid 21 and name HA) (resid 21 and name HD*) 2.0 0.2 2.5
assign (resid 21 and name HB*) (resid 21 and name HN) 2.0 0.2 3.5
assign (resid 21 and name HA) (resid 22 and name HN) 2.0 0.2 3.0
assign (resid 21 and name HN) (resid 22 and name HN) 2.0 0.2 1.5
assign (resid 21 and name HA) (resid 24 and name HN) 2.0 0.2 1.5
assign (resid 22 and name HB*) (resid 22 and name HN) 2.0 0.2 2.5
assign (resid 22 and name HA) (resid 23 and name HN) 2.0 0.2 1.5
assign (resid 22 and name HB*) (resid 23 and name HN) 2.0 0.2 3.0
assign (resid 22 and name HN) (resid 23 and name HN) 2.0 0.2 1.5
assign (resid 22 and name HA) (resid 25 and name HA1) 2.0 0.2 3
assign (resid 22 and name HA) (resid 25 and name HN) 2.0 0.2 2.5
assign (resid 22 and name HA) (resid 108 and name HG1*) 2.0 0.2 4.0
assign (resid 22 and name HB*) (resid 108 and name HA) 2.0 0.2 4.5
assign (resid 22 and name HB*) (resid 108 and name HD*) 2.0 0.2 2.5
assign (resid 22 and name HB*) (resid 108 and name HG2*) 2.0 0.2 3.2
assign (resid 23 and name HA) (resid 23 and name HD*) 2.0 0.2 3.0
assign (resid 23 and name HA) (resid 23 and name HD*) 2.0 0.2 4.4
assign (resid 23 and name HB*) (resid 23 and name HN) 2.0 0.2 2.5
assign (resid 23 and name HD1*) (resid 23 and name HN) 2.0 0.2 3.0
assign (resid 23 and name HD2*) (resid 23 and name HN) 2.0 0.2 3.0
assign (resid 23 and name HB*) (resid 24 and name HN) 2.0 0.2 2.5
assign (resid 23 and name HD1*) (resid 24 and name HN) 2.0 0.2 4.5
assign (resid 23 and name HD2*) (resid 24 and name HN) 2.0 0.2 4.5
assign (resid 23 and name HN) (resid 24 and name HN) 2.0 0.2 1.5
assign (resid 23 and name HN) (resid 25 and name HN) 2.0 0.2 2.5
assign (resid 23 and name HA) (resid 26 and name HN) 2.0 0.2 1.5
assign (resid 23 and name HA) (resid 28 and name HB*) 2.0 0.2 4.0
assign (resid 23 and name HA) (resid 28 and name HD*) 2.0 0.2 4.4
assign (resid 23 and name HD*) (resid 28 and name HB*) 2.0 0.2 5.4
assign (resid 23 and name HD*) (resid 30 and name HB*) 2.0 0.2 6.9
assign (resid 23 and name HD*) (resid 30 and name HB*) 2.0 0.2 7.0
assign (resid 23 and name HD*) (resid 108 and name HB) 2.0 0.2 5.9
assign (resid 23 and name HD*) (resid 108 and name HD1*) 2.0 0.2 4.6
assign (resid 24 and name HB*) (resid 24 and name HN) 2.0 0.2 2.5
assign (resid 24 and name HB*) (resid 25 and name HN) 2.0 0.2 3.0
assign (resid 24 and name HN) (resid 25 and name HN) 2.0 0.2 1.5
assign (resid 24 and name HN) (resid 26 and name HN) 2.0 0.2 2.5
assign (resid 24 and name HA) (resid 27 and name HN) 2.0 0.2 3.0
assign (resid 24 and name HA) (resid 28 and name HN) 2.0 0.2 3.0
assign (resid 25 and name HA*) (resid 25 and name HN) 2.0 0.2 1.5
assign (resid 25 and name HA*) (resid 26 and name HN) 2.0 0.2 2.5
assign (resid 25 and name HN) (resid 26 and name HN) 2.0 0.2 1.5
assign (resid 25 and name HA*) (resid 27 and name HN) 2.0 0.2 2.5
assign (resid 26 and name HB*) (resid 26 and name HN) 2.0 0.2 2.5
assign (resid 26 and name HA) (resid 27 and name HN) 2.0 0.2 3.0

assign (resid 26 and name HB*) (resid 27 and name HN) 2.0 0.2 3.0
assign (resid 26 and name HN) (resid 27 and name HN) 2.0 0.2 1.5
assign (resid 26 and name HB*) (resid 28 and name HD*) 2.0 0.2 4.6
assign (resid 26 and name HN) (resid 28 and name HN) 2.0 0.2 3.0
assign (resid 26 and name HB*) (resid 108 and name HB) 2.0 0.2 4.5
assign (resid 27 and name HA*) (resid 27 and name HN) 2.0 0.2 1.337
assign (resid 27 and name HA1) (resid 28 and name HN) 2.0 0.2 3.0
assign (resid 27 and name HA2) (resid 28 and name HN) 2.0 0.2 3.0
assign (resid 27 and name HN) (resid 28 and name HN) 2.0 0.2 1.5
assign (resid 28 and name HA) (resid 28 and name HD*) 2.0 0.2 3.6
assign (resid 28 and name HA) (resid 28 and name HG) 2.0 0.2 1.7
assign (resid 28 and name HA) (resid 29 and name HN) 2.0 0.2 0.5
assign (resid 28 and name HB1) (resid 29 and name HN) 2.0 0.2 1.5
assign (resid 28 and name HB2) (resid 29 and name HN) 2.0 0.2 1.5
assign (resid 28 and name HD1*) (resid 29 and name HN) 2.0 0.2 4.5
assign (resid 28 and name HD2*) (resid 29 and name HN) 2.0 0.2 4.5
assign (resid 28 and name HD*) (resid 112 and name HA) 2.0 0.2 4.2
assign (resid 28 and name HD*) (resid 112 and name HB*) 2.0 0.2 5.5
assign (resid 28 and name HD*) (resid 116 and name HG*) 2.0 0.2 5
assign (resid 28 and name HG*) (resid 116 and name HB*) 2.0 0.2 5
assign (resid 29 and name HA) (resid 29 and name HN) 2.0 0.2 1.5
assign (resid 29 and name HB) (resid 29 and name HN) 2.0 0.2 0.5
assign (resid 29 and name HA) (resid 30 and name HN) 2.0 0.2 0.5
assign (resid 29 and name HG2*) (resid 30 and name HN) 2.0 0.2 3.0
assign (resid 29 and name HB) (resid 112 and name HG*) 2.0 0.2 5.9
assign (resid 30 and name HB1) (resid 30 and name HN) 2.0 0.2 0.5
assign (resid 30 and name HB2) (resid 30 and name HN) 2.0 0.2 0.5
assign (resid 30 and name HA) (resid 31 and name HN) 2.0 0.2 0.5
assign (resid 30 and name HG) (resid 31 and name HN) 2.0 0.2 3.0
assign (resid 31 and name HB) (resid 31 and name HN) 2.0 0.2 1.5
assign (resid 31 and name HA) (resid 32 and name HN) 2.0 0.2 0.5
assign (resid 31 and name HG2*) (resid 32 and name HN) 2.0 0.2 1.5
assign (resid 31 and name HG2*) (resid 33 and name HE*) 2.0 0.2 5.0
assign (resid 31 and name HG2*) (resid 33 and name HZ) 2.0 0.2 3.0
assign (resid 32 and name HA) (resid 33 and name HD*) 2.0 0.2 5
assign (resid 32 and name HA) (resid 33 and name HN) 2.0 0.2 0.5
assign (resid 32 and name HB) (resid 39 and name HG*) 2.0 0.2 5.4
assign (resid 33 and name HB*) (resid 33 and name HN) 2.0 0.2 3.0
assign (resid 33 and name HA) (resid 34 and name HN) 2.0 0.2 1.5
assign (resid 33 and name HB1) (resid 35 and name HN) 2.0 0.2 3.0
assign (resid 33 and name HB2) (resid 35 and name HN) 2.0 0.2 3.0
assign (resid 33 and name HB*) (resid 38 and name HB*) 2.0 0.2 5.0
assign (resid 33 and name HD*) (resid 38 and name HB*) 2.0 0.2 4.5
assign (resid 33 and name HD*) (resid 38 and name HG*) 2.0 0.2 6.0
assign (resid 33 and name HB1) (resid 39 and name HG1*) 2.0 0.2 4.5
assign (resid 33 and name HB1) (resid 39 and name HG2*) 2.0 0.2 4.5
assign (resid 33 and name HB2) (resid 39 and name HG*) 2.0 0.2 5.9
assign (resid 33 and name HD*) (resid 39 and name HG*) 2.0 0.2 6.4
assign (resid 33 and name HD*) (resid 42 and name HB*) 2.0 0.2 5
assign (resid 33 and name HE*) (resid 42 and name HB*) 2.0 0.2 4.2
assign (resid 33 and name HZ) (resid 42 and name HB*) 2.0 0.2 2.7
assign (resid 33 and name HB*) (resid 51 and name HD*) 2.0 0.2 6.0
assign (resid 34 and name HA) (resid 34 and name HN) 2.0 0.2 1.5
assign (resid 34 and name HB*) (resid 34 and name HN) 2.0 0.2 1.7

assign (resid 34 and name HG*) (resid 34 and name HN) 2.0 0.2 3.5
assign (resid 34 and name HB*) (resid 35 and name HN) 2.0 0.2 4.0
assign (resid 34 and name HB*) (resid 35 and name HN) 2.0 0.2 4.5
assign (resid 34 and name HN) (resid 35 and name HN) 2.0 0.2 1.5
assign (resid 35 and name HD21) (resid 35 and name HB*) 2.0 0.2 2.7
assign (resid 35 and name HB*) (resid 36 and name HN) 2.0 0.2 2.5
assign (resid 35 and name HB*) (resid 37 and name HN) 2.0 0.2 4.0
assign (resid 35 and name HN) (resid 38 and name HB*) 2.0 0.2 2.5
assign (resid 36 and name HA*) (resid 36 and name HN) 2.0 0.2 2.5
assign (resid 36 and name HA*) (resid 37 and name HN) 2.0 0.2 2.5
assign (resid 36 and name HN) (resid 37 and name HN) 2.0 0.2 1.5
assign (resid 36 and name HA2) (resid 39 and name HB) 2.0 0.2 3.0
assign (resid 36 and name HA2) (resid 39 and name HG*) 2.0 0.2 5.9
assign (resid 36 and name HA2) (resid 39 and name HN) 2.0 0.2 3.5
assign (resid 36 and name HA*) (resid 40 and name HB*) 2.0 0.2 5
assign (resid 36 and name HA1) (resid 40 and name HD*) 2.0 0.2 5.9
assign (resid 36 and name HA2) (resid 40 and name HD*) 2.0 0.2 5.9
assign (resid 36 and name HA1) (resid 65 and name HD*) 2.0 0.2 4.4
assign (resid 36 and name HA2) (resid 65 and name HD*) 2.0 0.2 4.4
assign (resid 37 and name HA) (resid 37 and name HN) 2.0 0.2 1.5
assign (resid 37 and name HB1) (resid 37 and name HN) 2.0 0.2 1.5
assign (resid 37 and name HB1) (resid 38 and name HN) 2.0 0.2 1.5
assign (resid 37 and name HN) (resid 38 and name HN) 2.0 0.2 1.5
assign (resid 37 and name HN) (resid 39 and name HN) 2.0 0.2 2.5
assign (resid 37 and name HA) (resid 40 and name HB1) 2.0 0.2 3.0
assign (resid 37 and name HA) (resid 40 and name HB2) 2.0 0.2 3.0
assign (resid 37 and name HA) (resid 40 and name HN) 2.0 0.2 2.5
assign (resid 37 and name HA) (resid 41 and name HB*) 2.0 0.2 4.5
assign (resid 37 and name HA) (resid 65 and name HD*) 2.0 0.2 3.6
assign (resid 37 and name HA) (resid 68 and name HB*) 2.0 0.2 2.5
assign (resid 37 and name HA) (resid 68 and name HG*) 2.0 0.2 1.5
assign (resid 38 and name HA) (resid 38 and name HG1) 2.0 0.2 1.7
assign (resid 38 and name HA) (resid 38 and name HG2) 2.0 0.2 1.7
assign (resid 38 and name HA) (resid 38 and name HN) 2.0 0.2 1.5
assign (resid 38 and name HB*) (resid 39 and name HN) 2.0 0.2 3.5
assign (resid 38 and name HN) (resid 39 and name HN) 2.0 0.2 1.5
assign (resid 38 and name HN) (resid 41 and name HB*) 2.0 0.2 3.0
assign (resid 39 and name HB) (resid 39 and name HN) 2.0 0.2 0.5
assign (resid 39 and name HG1*) (resid 39 and name HA) 2.0 0.2 2.2
assign (resid 39 and name HG1*) (resid 39 and name HN) 2.0 0.2 2.2
assign (resid 39 and name HG2*) (resid 39 and name HA) 2.0 0.2 2.2
assign (resid 39 and name HG2*) (resid 39 and name HN) 2.0 0.2 2.2
assign (resid 39 and name HN) (resid 40 and name HN) 2.0 0.2 1.5
assign (resid 39 and name HA) (resid 41 and name HN) 2.0 0.2 2.5
assign (resid 39 and name HN) (resid 41 and name HN) 2.0 0.2 2.5
assign (resid 39 and name HA) (resid 42 and name HB*) 2.0 0.2 2.2
assign (resid 39 and name HA) (resid 42 and name HN) 2.0 0.2 1.5
assign (resid 39 and name HG*) (resid 42 and name HB*) 2.0 0.2 6.9
assign (resid 39 and name HA) (resid 43 and name HN) 2.0 0.2 2.5
assign (resid 39 and name HB) (resid 65 and name HD*) 2.0 0.2 5.9
assign (resid 39 and name HG*) (resid 68 and name HG*) 2.0 0.2 5.4
assign (resid 40 and name HA) (resid 41 and name HN) 2.0 0.2 2.5
assign (resid 40 and name HD*) (resid 41 and name HA) 2.0 0.2 6.0
assign (resid 40 and name HN) (resid 41 and name HN) 2.0 0.2 1.5

assign (resid 40 and name HA) (resid 43 and name HB*) 2.0 0.2 4
assign (resid 40 and name HA) (resid 43 and name HD*) 2.0 0.2 6.4
assign (resid 40 and name HA) (resid 43 and name HN) 2.0 0.2 1.5
assign (resid 40 and name HA) (resid 44 and name HN) 2.0 0.2 2.5
assign (resid 40 and name HD1*) (resid 68 and name HG*) 2.0 0.2 5.5
assign (resid 40 and name HD2*) (resid 68 and name HG*) 2.0 0.2 5.5
assign (resid 40 and name HD*) (resid 69 and name HA) 2.0 0.2 5.9
assign (resid 40 and name HG) (resid 69 and name HB) 2.0 0.2 3.0
assign (resid 41 and name HB*) (resid 41 and name HN) 2.0 0.2 2.5
assign (resid 41 and name HN) (resid 43 and name HN) 2.0 0.2 2.5
assign (resid 41 and name HA) (resid 44 and name HB*) 2.0 0.2 3.0
assign (resid 42 and name HA) (resid 43 and name HN) 2.0 0.2 1.5
assign (resid 42 and name HB*) (resid 43 and name HN) 2.0 0.2 3.0
assign (resid 42 and name HN) (resid 43 and name HN) 2.0 0.2 1.5
assign (resid 43 and name HA) (resid 43 and name HN) 2.0 0.2 1.5
assign (resid 43 and name HD*) (resid 43 and name HA) 2.0 0.2 4.4
assign (resid 43 and name HD*) (resid 43 and name HA) 2.0 0.2 4.4
assign (resid 43 and name HG) (resid 43 and name HA) 2.0 0.2 3
assign (resid 43 and name HB*) (resid 44 and name HN) 2.0 0.2 1.5
assign (resid 43 and name HB*) (resid 44 and name HN) 2.0 0.2 3.0
assign (resid 43 and name HN) (resid 44 and name HN) 2.0 0.2 1.5
assign (resid 43 and name HA) (resid 45 and name HN) 2.0 0.2 3.0
assign (resid 43 and name HA) (resid 46 and name HN) 2.0 0.2 1.5
assign (resid 43 and name HD*) (resid 46 and name HB*) 2.0 0.2 5.5
assign (resid 43 and name HD*) (resid 47 and name HA) 2.0 0.2 4.4
assign (resid 43 and name HD*) (resid 69 and name HA) 2.0 0.2 5.9
assign (resid 43 and name HA) (resid 73 and name HD2) 2.0 0.2 3
assign (resid 43 and name HD1*) (resid 73 and name HD2) 2.0 0.2 3.0
assign (resid 43 and name HD2*) (resid 73 and name HD2) 2.0 0.2 3.0
assign (resid 43 and name HG) (resid 73 and name HD2) 2.0 0.2 3
assign (resid 43 and name HN) (resid 73 and name HD2) 2.0 0.2 3.0
assign (resid 43 and name HD*) (resid 74 and name HD*) 2.0 0.2 6.9
assign (resid 44 and name HB*) (resid 44 and name HN) 2.0 0.2 2.5
assign (resid 44 and name HA) (resid 45 and name HN) 2.0 0.2 1.5
assign (resid 44 and name HB*) (resid 45 and name HB*) 2.0 0.2 5.5
assign (resid 44 and name HB*) (resid 45 and name HN) 2.0 0.2 3.5
assign (resid 44 and name HN) (resid 45 and name HN) 2.0 0.2 1.5
assign (resid 44 and name HA) (resid 46 and name HN) 2.0 0.2 3.0
assign (resid 44 and name HN) (resid 46 and name HN) 2.0 0.2 3.0
assign (resid 45 and name HB*) (resid 46 and name HN) 2.0 0.2 4
assign (resid 45 and name HN) (resid 46 and name HN) 2.0 0.2 0.7
assign (resid 46 and name HD*) (resid 46 and name HA) 2.0 0.2 1.7
assign (resid 46 and name HD*) (resid 46 and name HB*) 2.0 0.2 2.1
assign (resid 46 and name HD1*) (resid 46 and name HA) 2.0 0.2 3.0
assign (resid 46 and name HD2*) (resid 46 and name HA) 2.0 0.2 3.0
assign (resid 46 and name HG*) (resid 46 and name HN) 2.0 0.2 2.5
assign (resid 46 and name HA) (resid 47 and name HN) 2.0 0.2 0.7
assign (resid 46 and name HB*) (resid 47 and name HN) 2.0 0.2 2.5
assign (resid 46 and name HG*) (resid 47 and name HN) 2.0 0.2 3.5
assign (resid 46 and name HN) (resid 47 and name HN) 2.0 0.2 3.0
assign (resid 46 and name HB*) (resid 74 and name HB*) 2.0 0.2 5.5
assign (resid 46 and name HE*) (resid 74 and name HA) 2.0 0.2 4
assign (resid 46 and name HG*) (resid 76 and name HD*) 2.0 0.2 6.9
assign (resid 47 and name HB) (resid 47 and name HN) 2.0 0.2 1.5

assign (resid 47 and name HG2*) (resid 47 and name HN) 2.0 0.2 3.0
assign (resid 47 and name HA) (resid 51 and name HD*) 2.0 0.2 5.9
assign (resid 47 and name HA) (resid 76 and name HD*) 2.0 0.2 5.9
assign (resid 47 and name HB) (resid 76 and name HD*) 2.0 0.2 5.9
assign (resid 49 and name HB*) (resid 49 and name HN) 2.0 0.2 2.5
assign (resid 49 and name HA) (resid 50 and name HG*) 2.0 0.2 5.9
assign (resid 49 and name HA) (resid 50 and name HN) 2.0 0.2 2.5
assign (resid 49 and name HB*) (resid 50 and name HN) 2.0 0.2 2.5
assign (resid 49 and name HN) (resid 50 and name HN) 2.0 0.2 0.7
assign (resid 49 and name HA) (resid 76 and name HD*) 2.0 0.2 5.9
assign (resid 49 and name HA) (resid 119 and name HD*) 2.0 0.2 4.5
assign (resid 49 and name HB*) (resid 119 and name HG2*) 2.0 0.2 5.5
assign (resid 50 and name HA) (resid 50 and name HN) 2.0 0.2 1.5
assign (resid 50 and name HB) (resid 50 and name HN) 2.0 0.2 1.5
assign (resid 50 and name HG*) (resid 50 and name HA) 2.0 0.2 3.6
assign (resid 50 and name HN) (resid 50 and name HG*) 2.0 0.2 3.6
assign (resid 50 and name HA) (resid 51 and name HN) 2.0 0.2 0.7
assign (resid 50 and name HG*) (resid 51 and name HN) 2.0 0.2 3.6
assign (resid 50 and name HG*) (resid 52 and name HG) 2.0 0.2 3.6
assign (resid 50 and name HA) (resid 76 and name HD*) 2.0 0.2 3.4
assign (resid 50 and name HN) (resid 76 and name HD*) 2.0 0.2 5.9
assign (resid 50 and name HA) (resid 77 and name HD*) 2.0 0.2 4.0
assign (resid 50 and name HA) (resid 77 and name HG1) 2.0 0.2 3.0
assign (resid 50 and name HG1*) (resid 77 and name HG*) 2.0 0.2 3.5
assign (resid 50 and name HG2*) (resid 77 and name HG*) 2.0 0.2 3.5
assign (resid 50 and name HB) (resid 119 and name HD*) 2.0 0.2 5.9
assign (resid 50 and name HB) (resid 119 and name HG2*) 2.0 0.2 4.2
assign (resid 50 and name HG*) (resid 119 and name HA) 2.0 0.2 4.4
assign (resid 50 and name HG*) (resid 119 and name HD*) 2.0 0.2 4.6
assign (resid 50 and name HG1*) (resid 119 and name HG1*) 2.0 0.2 3.5
assign (resid 50 and name HG2*) (resid 119 and name HG1*) 2.0 0.2 3.5
assign (resid 51 and name HB*) (resid 51 and name HN) 2.0 0.2 2.5
assign (resid 51 and name HD1*) (resid 51 and name HA) 2.0 0.2 3.0
assign (resid 51 and name HD2*) (resid 51 and name HA) 2.0 0.2 3.0
assign (resid 51 and name HN) (resid 51 and name HD*) 2.0 0.2 4.4
assign (resid 51 and name HA) (resid 52 and name HN) 2.0 0.2 0.5
assign (resid 51 and name HG) (resid 69 and name HG1*) 2.0 0.2 2.5
assign (resid 51 and name HB*) (resid 76 and name HD*) 2.0 0.2 6.9
assign (resid 51 and name HN) (resid 76 and name HD*) 2.0 0.2 5.9
assign (resid 51 and name HB1) (resid 78 and name HA) 2.0 0.2 3.0
assign (resid 51 and name HB2) (resid 78 and name HA) 2.0 0.2 3.0
assign (resid 51 and name HD*) (resid 78 and name HA) 2.0 0.2 5.9
assign (resid 51 and name HN) (resid 78 and name HA) 2.0 0.2 2.5
assign (resid 51 and name HB*) (resid 99 and name HE*) 2.0 0.2 6.0
assign (resid 52 and name HB*) (resid 52 and name HN) 2.0 0.2 2.5
assign (resid 52 and name HD*) (resid 52 and name HA) 2.0 0.2 3.6
assign (resid 52 and name HA) (resid 53 and name HN) 2.0 0.2 1.5
assign (resid 52 and name HD*) (resid 54 and name HA) 2.0 0.2 5.9
assign (resid 52 and name HA) (resid 79 and name HB) 2.0 0.2 1.5
assign (resid 52 and name HA) (resid 79 and name HG2*) 2.0 0.2 4.2
assign (resid 52 and name HA) (resid 79 and name HN) 2.0 0.2 2.5
assign (resid 52 and name HB*) (resid 79 and name HG2*) 2.0 0.2 5.5
assign (resid 52 and name HD*) (resid 79 and name HB) 2.0 0.2 4.4
assign (resid 52 and name HD*) (resid 79 and name HG2*) 2.0 0.2 6.6

assign (resid 52 and name HG) (resid 79 and name HG2*) 2.0 0.2 2.2
assign (resid 52 and name HD*) (resid 81 and name HE*) 2.0 0.2 5.4
assign (resid 52 and name HB*) (resid 81 and name HE*) 2.0 0.2 5.5
assign (resid 53 and name HA) (resid 54 and name HN) 2.0 0.2 0.5
assign (resid 53 and name HN) (resid 81 and name HN) 2.0 0.2 3.0
assign (resid 54 and name HB*) (resid 54 and name HN) 2.0 0.2 3.5
assign (resid 54 and name HN) (resid 54 and name HB*) 2.0 0.2 1.7
assign (resid 54 and name HB*) (resid 63 and name HD*) 2.0 0.2 5.4
assign (resid 54 and name HA) (resid 81 and name HN) 2.0 0.2 2.5
assign (resid 55 and name HA) (resid 55 and name HD*) 2.0 0.2 2.2
assign (resid 55 and name HA) (resid 62 and name HA1) 2.0 0.2 3.0
assign (resid 55 and name HA) (resid 62 and name HA2) 2.0 0.2 3.0
assign (resid 55 and name HG2*) (resid 82 and name HA) 2.0 0.2 3.0
assign (resid 55 and name HG1*) (resid 84 and name HB1) 2.0 0.2 4
assign (resid 55 and name HG1*) (resid 84 and name HB2) 2.0 0.2 4
assign (resid 57 and name HA) (resid 57 and name HG1) 2.0 0.2 3.0
assign (resid 57 and name HA) (resid 57 and name HG2) 2.0 0.2 3.0
assign (resid 57 and name HA) (resid 58 and name HD*) 2.0 0.2 2.5
assign (resid 57 and name HA) (resid 58 and name HD1) 2.0 0.2 1.7
assign (resid 57 and name HA) (resid 58 and name HD2) 2.0 0.2 1.7
assign (resid 57 and name HB*) (resid 58 and name HD1) 2.0 0.2 4.0
assign (resid 57 and name HB*) (resid 58 and name HD2) 2.0 0.2 4.0
assign (resid 57 and name HE*) (resid 65 and name HD*) 2.0 0.2 4.6
assign (resid 58 and name HA) (resid 59 and name HN) 2.0 0.2 0.5
assign (resid 58 and name HB*) (resid 59 and name HN) 2.0 0.2 6.0
assign (resid 62 and name HA*) (resid 63 and name HN) 2.0 0.2 2.5
assign (resid 62 and name HN) (resid 63 and name HN) 2.0 0.2 1.5
assign (resid 62 and name HA*) (resid 65 and name HD*) 2.0 0.2 6.9
assign (resid 63 and name HA) (resid 63 and name HD*) 2.0 0.2 3.6
assign (resid 63 and name HA) (resid 64 and name HN) 2.0 0.2 1.5
assign (resid 63 and name HN) (resid 64 and name HN) 2.0 0.2 3.0
assign (resid 65 and name HA) (resid 68 and name HG*) 2.0 0.2 2.5
assign (resid 65 and name HD*) (resid 68 and name HG*) 2.0 0.2 5.4
assign (resid 66 and name HA) (resid 66 and name HD*) 2.0 0.2 5.6
assign (resid 66 and name HA) (resid 69 and name HB) 2.0 0.2 1.5
assign (resid 66 and name HA) (resid 69 and name HD1*) 2.0 0.2 3.0
assign (resid 66 and name HA) (resid 69 and name HG2*) 2.0 0.2 6.5
assign (resid 66 and name HA) (resid 69 and name HN) 2.0 0.2 1.5
assign (resid 66 and name HD*) (resid 78 and name HB) 2.0 0.2 5.9
assign (resid 66 and name HD*) (resid 99 and name HZ) 2.0 0.2 5.9
assign (resid 67 and name HB*) (resid 67 and name HD*) 2.0 0.2 2.7
assign (resid 67 and name HB*) (resid 67 and name HG1) 2.0 0.2 1.7
assign (resid 67 and name HB*) (resid 67 and name HG2) 2.0 0.2 1.7
assign (resid 67 and name HB1) (resid 67 and name HE*) 2.0 0.2 4.0
assign (resid 67 and name HB2) (resid 67 and name HE*) 2.0 0.2 4.0
assign (resid 67 and name HD*) (resid 67 and name HA) 2.0 0.2 2.5
assign (resid 67 and name HA) (resid 70 and name HB*) 2.0 0.2 2.5
assign (resid 67 and name HB*) (resid 89 and name HA) 2.0 0.2 2.5
assign (resid 67 and name HD*) (resid 89 and name HA) 2.0 0.2 2.5
assign (resid 68 and name HA) (resid 69 and name HN) 2.0 0.2 3.0
assign (resid 68 and name HB*) (resid 69 and name HN) 2.0 0.2 2.5
assign (resid 68 and name HN) (resid 69 and name HN) 2.0 0.2 1.5
assign (resid 69 and name HB) (resid 69 and name HN) 2.0 0.2 0.5
assign (resid 69 and name HG*) (resid 69 and name HA) 2.0 0.2 2.5

assign (resid 69 and name HN) (resid 69 and name HD*) 2.0 0.2 3.0
assign (resid 69 and name HN) (resid 69 and name HG2*) 2.0 0.2 3.0
assign (resid 69 and name HB*) (resid 70 and name HN) 2.0 0.2 2.5
assign (resid 69 and name HG2*) (resid 70 and name HN) 2.0 0.2 4.5
assign (resid 69 and name HN) (resid 70 and name HN) 2.0 0.2 1.5
assign (resid 69 and name HA) (resid 72 and name HB1) 2.0 0.2 3
assign (resid 69 and name HA) (resid 72 and name HB2) 2.0 0.2 3
assign (resid 69 and name HA) (resid 73 and name HD2) 2.0 0.2 3.0
assign (resid 69 and name HA) (resid 73 and name HN) 2.0 0.2 3.0
assign (resid 70 and name HA) (resid 70 and name HD*) 2.0 0.2 1.7
assign (resid 70 and name HA) (resid 70 and name HG*) 2.0 0.2 1.7
assign (resid 70 and name HN) (resid 70 and name HD*) 2.0 0.2 4.0
assign (resid 70 and name HN) (resid 71 and name HN) 2.0 0.2 3.0
assign (resid 70 and name HB*) (resid 72 and name HN) 2.0 0.2 4.0
assign (resid 70 and name HA) (resid 73 and name HN) 2.0 0.2 2.5
assign (resid 70 and name HA) (resid 78 and name HG*) 2.0 0.2 5.9
assign (resid 70 and name HG*) (resid 78 and name HG1*) 2.0 0.2 5.5
assign (resid 70 and name HG*) (resid 78 and name HG2*) 2.0 0.2 5.5
assign (resid 71 and name HA) (resid 71 and name HG*) 2.0 0.2 1.7
assign (resid 71 and name HB*) (resid 72 and name HD1) 2.0 0.2 4.0
assign (resid 71 and name HB*) (resid 72 and name HD2) 2.0 0.2 4.0
assign (resid 71 and name HN) (resid 72 and name HN) 2.0 0.2 1.5
assign (resid 72 and name HA) (resid 72 and name HD*) 2.0 0.2 2.5
assign (resid 72 and name HD*) (resid 72 and name HB*) 2.0 0.2 2.7
assign (resid 72 and name HB*) (resid 73 and name HN) 2.0 0.2 4.0
assign (resid 72 and name HN) (resid 73 and name HN) 2.0 0.2 1.5
assign (resid 73 and name HA) (resid 74 and name HD*) 2.0 0.2 1.7
assign (resid 73 and name HA) (resid 74 and name HD1) 2.0 0.2 0.7
assign (resid 73 and name HA) (resid 74 and name HD2) 2.0 0.2 0.7
assign (resid 73 and name HB*) (resid 76 and name HD*) 2.0 0.2 6.9
assign (resid 73 and name HB*) (resid 76 and name HN) 2.0 0.2 5.4
assign (resid 73 and name HD2) (resid 76 and name HD1*) 2.0 0.2 4.5
assign (resid 73 and name HD2) (resid 76 and name HD2*) 2.0 0.2 4.5
assign (resid 74 and name HD*) (resid 75 and name HN) 2.0 0.2 4.0
assign (resid 75 and name HA) (resid 75 and name HG*) 2.0 0.2 2.1
assign (resid 75 and name HA) (resid 75 and name HG*) 2.0 0.2 2.5
assign (resid 75 and name HG*) (resid 75 and name HN) 2.0 0.2 1.5
assign (resid 75 and name HN) (resid 76 and name HN) 2.0 0.2 1.5
assign (resid 76 and name HD*) (resid 76 and name HA) 2.0 0.2 3.6
assign (resid 76 and name HD*) (resid 76 and name HA) 2.0 0.2 4.4
assign (resid 76 and name HN) (resid 76 and name HB*) 2.0 0.2 1.7
assign (resid 76 and name HN) (resid 76 and name HD*) 2.0 0.2 5.9
assign (resid 76 and name HN) (resid 76 and name HG) 2.0 0.2 0.5
assign (resid 76 and name HA) (resid 77 and name HD*) 2.0 0.2 2.5
assign (resid 76 and name HA) (resid 77 and name HD1) 2.0 0.2 1.7
assign (resid 76 and name HA) (resid 77 and name HD2) 2.0 0.2 1.7
assign (resid 76 and name HD*) (resid 77 and name HD*) 2.0 0.2 6.9
assign (resid 76 and name HD*) (resid 77 and name HG*) 2.0 0.2 6.9
assign (resid 76 and name HD*) (resid 77 and name HD*) 2.0 0.2 6.9
assign (resid 76 and name HA) (resid 78 and name HG*) 2.0 0.2 5.9
assign (resid 76 and name HD*) (resid 78 and name HG*) 2.0 0.2 6.8
assign (resid 76 and name HG) (resid 78 and name HG*) 2.0 0.2 5.91554
assign (resid 76 and name HG) (resid 78 and name HG1*) 2.0 0.2 4.5
assign (resid 76 and name HG) (resid 78 and name HG2*) 2.0 0.2 4.5

assign (resid 76 and name HD*) (resid 119 and name HD1*) 2.0 0.2 5.5
assign (resid 76 and name HD*) (resid 119 and name HG2*) 2.0 0.2 4.6
assign (resid 77 and name HA) (resid 78 and name HG*) 2.0 0.2 5.9
assign (resid 77 and name HA) (resid 78 and name HN) 2.0 0.2 0.5
assign (resid 77 and name HG*) (resid 98 and name HB*) 2.0 0.2 5.5
assign (resid 77 and name HG*) (resid 122 and name HE*) 2.0 0.2 6.0
assign (resid 78 and name HA) (resid 78 and name HG*) 2.0 0.2 3.6
assign (resid 78 and name HN) (resid 78 and name HB) 2.0 0.2 1.5
assign (resid 78 and name HN) (resid 78 and name HG1*) 2.0 0.2 4.5
assign (resid 78 and name HN) (resid 78 and name HG2*) 2.0 0.2 4.5
assign (resid 78 and name HA) (resid 79 and name HN) 2.0 0.2 0.5
assign (resid 78 and name HG*) (resid 99 and name HZ) 2.0 0.2 5.9
assign (resid 79 and name HD1*) (resid 115 and name HG*) 2.0 0.2 4.6
assign (resid 81 and name HA) (resid 81 and name HG1) 2.0 0.2 3.0
assign (resid 81 and name HA) (resid 81 and name HG2) 2.0 0.2 3.0
assign (resid 81 and name HA) (resid 82 and name HN) 2.0 0.2 0.5
assign (resid 81 and name HA) (resid 102 and name HB*) 2.0 0.2 3.5
assign (resid 81 and name HA) (resid 102 and name HD*) 2.0 0.2 6.0
assign (resid 81 and name HA) (resid 102 and name HN) 2.0 0.2 3.0
assign (resid 81 and name HB*) (resid 102 and name HD*) 2.0 0.2 5.5
assign (resid 81 and name HG*) (resid 102 and name HD*) 2.0 0.2 4.5
assign (resid 81 and name HG*) (resid 104 and name HG*) 2.0 0.2 5.0
assign (resid 82 and name HN) (resid 82 and name HG2*) 2.0 0.2 4.4
assign (resid 82 and name HG2*) (resid 101 and name HD*) 2.0 0.2 6.5
assign (resid 82 and name HG2*) (resid 101 and name HE*) 2.0 0.2 6.5
assign (resid 82 and name HG1*) (resid 103 and name HA) 2.0 0.2 3.0
assign (resid 82 and name HN) (resid 103 and name HA) 2.0 0.2 1.5
assign (resid 84 and name HA) (resid 87 and name HB*) 2.0 0.2 4.0
assign (resid 85 and name HA) (resid 88 and name HB*) 2.0 0.2 4
assign (resid 86 and name HA) (resid 87 and name HN) 2.0 0.2 3.0
assign (resid 86 and name HB*) (resid 87 and name HN) 2.0 0.2 4.0
assign (resid 86 and name HN) (resid 87 and name HN) 2.0 0.2 3.0
assign (resid 87 and name HA) (resid 87 and name HD*) 2.0 0.2 3.6
assign (resid 87 and name HB*) (resid 88 and name HN) 2.0 0.2 2.5
assign (resid 87 and name HN) (resid 88 and name HN) 2.0 0.2 2.5
assign (resid 87 and name HA) (resid 90 and name HB*) 2.0 0.2 3.0
assign (resid 87 and name HD*) (resid 90 and name HB*) 2.0 0.2 4.6
assign (resid 87 and name HA) (resid 91 and name HG*) 2.0 0.2 4.4
assign (resid 87 and name HD*) (resid 101 and name HD*) 2.0 0.2 6.4
assign (resid 88 and name HN) (resid 88 and name HB*) 2.0 0.2 1.7
assign (resid 88 and name HB*) (resid 89 and name HN) 2.0 0.2 2.5
assign (resid 88 and name HN) (resid 89 and name HN) 2.0 0.2 1.5
assign (resid 88 and name HA) (resid 91 and name HB) 2.0 0.2 1.5
assign (resid 88 and name HA) (resid 91 and name HN) 2.0 0.2 2.5
assign (resid 89 and name HB*) (resid 90 and name HN) 2.0 0.2 3.0
assign (resid 89 and name HA) (resid 92 and name HB*) 2.0 0.2 2.5
assign (resid 89 and name HA) (resid 92 and name HN) 2.0 0.2 2.5
assign (resid 90 and name HB*) (resid 91 and name HN) 2.0 0.2 3.0
assign (resid 90 and name HN) (resid 91 and name HN) 2.0 0.2 1.5
assign (resid 90 and name HA) (resid 93 and name HB*) 2.0 0.2 4.5
assign (resid 90 and name HB*) (resid 101 and name HD*) 2.0 0.2 6.5
assign (resid 90 and name HB*) (resid 101 and name HE*) 2.0 0.2 6.5
assign (resid 91 and name HG1*) (resid 91 and name HA) 2.0 0.2 3.0
assign (resid 91 and name HG2*) (resid 91 and name HA) 2.0 0.2 3.0

assign (resid 91 and name HN) (resid 91 and name HB) 2.0 0.2 0.7
assign (resid 91 and name HN) (resid 91 and name HG1*) 2.0 0.2 4.5
assign (resid 91 and name HN) (resid 91 and name HG2*) 2.0 0.2 4.5
assign (resid 91 and name HB) (resid 92 and name HN) 2.0 0.2 1.5
assign (resid 91 and name HG1*) (resid 92 and name HN) 2.0 0.2 4.0
assign (resid 91 and name HG2*) (resid 92 and name HN) 2.0 0.2 4.0
assign (resid 91 and name HN) (resid 92 and name HN) 2.0 0.2 1.5
assign (resid 91 and name HG*) (resid 95 and name HB*) 2.0 0.2 5.4
assign (resid 91 and name HG*) (resid 101 and name HE*) 2.0 0.2 7.9
assign (resid 91 and name HA) (resid 101 and name HE*) 2.0 0.2 5.0
assign (resid 92 and name HN) (resid 93 and name HN) 2.0 0.2 3.0
assign (resid 92 and name HA) (resid 95 and name HB*) 2.0 0.2 2.5
assign (resid 92 and name HA) (resid 95 and name HG*) 2.0 0.2 2.5
assign (resid 97 and name HA*) (resid 98 and name HN) 2.0 0.2 2.5
assign (resid 98 and name HB*) (resid 99 and name HN) 2.0 0.2 4.0
assign (resid 98 and name HN) (resid 99 and name HN) 2.0 0.2 3.0
assign (resid 99 and name HN) (resid 99 and name HB1) 2.0 0.2 3.0
assign (resid 99 and name HN) (resid 99 and name HB2) 2.0 0.2 3.0
assign (resid 99 and name HN) (resid 100 and name HN) 2.0 0.2 3.0
assign (resid 100 and name HN) (resid 100 and name HA) 2.0 0.2 1.7
assign (resid 100 and name HN) (resid 100 and name HB1) 2.0 0.2 1.7
assign (resid 100 and name HN) (resid 100 and name HB2) 2.0 0.2 1.7
assign (resid 101 and name HA) (resid 102 and name HN) 2.0 0.2 0.5
assign (resid 101 and name HB*) (resid 102 and name HB*) 2.0 0.2 5
assign (resid 102 and name HB1) (resid 102 and name HD*) 2.0 0.2 4.4
assign (resid 102 and name HB2) (resid 102 and name HD*) 2.0 0.2 4.4
assign (resid 102 and name HD*) (resid 102 and name HA) 2.0 0.2 3.6
assign (resid 102 and name HA) (resid 103 and name HD*) 2.0 0.2 2.5
assign (resid 102 and name HD*) (resid 103 and name HD1) 2.0 0.2 4.4
assign (resid 102 and name HD*) (resid 103 and name HD2) 2.0 0.2 4.4
assign (resid 102 and name HB*) (resid 106 and name HD*) 2.0 0.2 6.0
assign (resid 102 and name HD*) (resid 106 and name HB1) 2.0 0.2 4.4
assign (resid 102 and name HD*) (resid 106 and name HB2) 2.0 0.2 4.4
assign (resid 102 and name HD1*) (resid 106 and name HD*) 2.0 0.2 6.5
assign (resid 102 and name HD2*) (resid 106 and name HD*) 2.0 0.2 6.5
assign (resid 102 and name HD*) (resid 111 and name HB*) 2.0 0.2 4.6
assign (resid 102 and name HD1*) (resid 111 and name HA) 2.0 0.2 3.0
assign (resid 102 and name HD2*) (resid 111 and name HA) 2.0 0.2 3.0
assign (resid 104 and name HA) (resid 105 and name HA) 2.0 0.2 0.5
assign (resid 104 and name HA) (resid 106 and name HD1) 2.0 0.2 3.0
assign (resid 104 and name HA) (resid 106 and name HN) 2.0 0.2 1.5
assign (resid 105 and name HA) (resid 106 and name HN) 2.0 0.2 0.5
assign (resid 105 and name HB*) (resid 106 and name HN) 2.0 0.2 4.0
assign (resid 106 and name HD*) (resid 106 and name HA) 2.0 0.2 5.0
assign (resid 106 and name HN) (resid 106 and name HD1) 2.0 0.2 3.0
assign (resid 106 and name HA) (resid 107 and name HN) 2.0 0.2 0.7
assign (resid 106 and name HB1) (resid 107 and name HN) 2.0 0.2 2.5
assign (resid 106 and name HB2) (resid 107 and name HN) 2.0 0.2 1.5
assign (resid 106 and name HD2) (resid 107 and name HN) 2.0 0.2 3.0
assign (resid 106 and name HD2) (resid 108 and name HD*) 2.0 0.2 4.5
assign (resid 106 and name HD2) (resid 108 and name HG1*) 2.0 0.2 4.0
assign (resid 106 and name HE2) (resid 108 and name HA) 2.0 0.2 1.5
assign (resid 107 and name HN) (resid 107 and name HB1) 2.0 0.2 0.7
assign (resid 107 and name HN) (resid 107 and name HB2) 2.0 0.2 0.7

assign (resid 107 and name HA) (resid 108 and name HN) 2.0 0.2 1.5
assign (resid 107 and name HB1) (resid 110 and name HB*) 2.0 0.2 2.5
assign (resid 107 and name HB2) (resid 110 and name HB*) 2.0 0.2 2.5
assign (resid 107 and name HB1) (resid 110 and name HG1) 2.0 0.2 3.0
assign (resid 107 and name HB1) (resid 110 and name HG2) 2.0 0.2 3.0
assign (resid 107 and name HB2) (resid 110 and name HG1) 2.0 0.2 3.0
assign (resid 107 and name HB2) (resid 110 and name HG2) 2.0 0.2 3.0
assign (resid 107 and name HN) (resid 110 and name HB1) 2.0 0.2 3.0
assign (resid 107 and name HN) (resid 110 and name HB2) 2.0 0.2 3.0
assign (resid 108 and name HA) (resid 108 and name HD*) 2.0 0.2 2.2
assign (resid 108 and name HB) (resid 108 and name HD*) 2.0 0.2 3.0
assign (resid 108 and name HG*) (resid 108 and name HA) 2.0 0.2 2.2
assign (resid 108 and name HG*) (resid 108 and name HA) 2.0 0.2 2.5
assign (resid 108 and name HN) (resid 108 and name HB) 2.0 0.2 0.5
assign (resid 108 and name HN) (resid 108 and name HG1*) 2.0 0.2 3.0
assign (resid 108 and name HG2*) (resid 109 and name HA) 2.0 0.2 4.5
assign (resid 108 and name HN) (resid 109 and name HN) 2.0 0.2 1.5
assign (resid 108 and name.HG1*) (resid 110 and name HN) 2.0 0.2 4.5
assign (resid 108 and name HA) (resid 111 and name HN) 2.0 0.2 1.5
assign (resid 108 and name HA) (resid 112 and name HG2*) 2.0 0.2 4.0
assign (resid 108 and name HA) (resid 112 and name HN) 2.0 0.2 2.5
assign (resid 108 and name HA) (resid 112 and name HN) 2.0 0.2 2.5
assign (resid 108 and name HG*) (resid 112 and name HG*) 2.0 0.2 5.4
assign (resid 109 and name HB1) (resid 109 and name HN) 2.0 0.2 0.5
assign (resid 109 and name HB2) (resid 109 and name HN) 2.0 0.2 0.5
assign (resid 109 and name HB1) (resid 110 and name HN) 2.0 0.2 1.5
assign (resid 109 and name HB2) (resid 110 and name HN) 2.0 0.2 1.5
assign (resid 109 and name HA) (resid 112 and name HN) 2.0 0.2 1.5
assign (resid 109 and name HA) (resid 113 and name HN) 2.0 0.2 1.5
assign (resid 110 and name HB1) (resid 110 and name HN) 2.0 0.2 0.5
assign (resid 110 and name HB2) (resid 110 and name HN) 2.0 0.2 0.5
assign (resid 110 and name HG*) (resid 110 and name HN) 2.0 0.2 2.5
assign (resid 110 and name HB1) (resid 111 and name HN) 2.0 0.2 5.0
assign (resid 110 and name HB2) (resid 111 and name HN) 2.0 0.2 5.0
assign (resid 110 and name HN) (resid 111 and name HN) 2.0 0.2 1.5
assign (resid 110 and name HA) (resid 113 and name HN) 2.0 0.2 3.0
assign (resid 110 and name HA) (resid 113 and name HB*) 2.0 0.2 2.2
assign (resid 110 and name HA) (resid 114 and name HB*) 2.0 0.2 4.0
assign (resid 111 and name HB*) (resid 111 and name HN) 2.0 0.2 0.5
assign (resid 111 and name HB*) (resid 112 and name HN) 2.0 0.2 3.0
assign (resid 111 and name HA) (resid 114 and name HB*) 2.0 0.2 2.5
assign (resid 111 and name HA) (resid 114 and name HD1*) 2.0 0.2 4.5
assign (resid 111 and name HA) (resid 114 and name HD2*) 2.0 0.2 4.5
assign (resid 111 and name HA) (resid 114 and name HN) 2.0 0.2 2.5
assign (resid 111 and name HA) (resid 115 and name HN) 2.0 0.2 2.5
assign (resid 112 and name HB) (resid 112 and name HN) 2.0 0.2 0.5
assign (resid 112 and name HG1*) (resid 112 and name HA) 2.0 0.2 2.2
assign (resid 112 and name HG1*) (resid 112 and name HN) 2.0 0.2 2.5
assign (resid 112 and name HG2*) (resid 112 and name HA) 2.0 0.2 2.2
assign (resid 112 and name HG2*) (resid 112 and name HN) 2.0 0.2 2.5
assign (resid 112 and name HB) (resid 113 and name HN) 2.0 0.2 1.5
assign (resid 112 and name HN) (resid 113 and name HN) 2.0 0.2 1.5
assign (resid 112 and name HA) (resid 115 and name HB) 2.0 0.2 1.7
assign (resid 113 and name HB*) (resid 113 and name HN) 2.0 0.2 2.2

assign (resid 113 and name HB*) (resid 114 and name HN) 2.0 0.2 3.0
assign (resid 113 and name HN) (resid 114 and name HN) 2.0 0.2 1.5
assign (resid 113 and name HA) (resid 116 and name HB1) 2.0 0.2 2.5
assign (resid 113 and name HA) (resid 116 and name HB2) 2.0 0.2 2.5
assign (resid 113 and name HA) (resid 116 and name HG*) 2.0 0.2 4.2
assign (resid 113 and name HA) (resid 116 and name HN) 2.0 0.2 1.5
assign (resid 113 and name HB*) (resid 116 and name HN) 2.0 0.2 4.5
assign (resid 114 and name HN) (resid 114 and name HB1) 2.0 0.2 0.5
assign (resid 114 and name HN) (resid 114 and name HB2) 2.0 0.2 0.5
assign (resid 114 and name HN) (resid 114 and name HD*) 2.0 0.2 5.9
assign (resid 114 and name HB*) (resid 115 and name HN) 2.0 0.2 2.5
assign (resid 114 and name HB1) (resid 115 and name HN) 2.0 0.2 1.7
assign (resid 114 and name HB2) (resid 115 and name HN) 2.0 0.2 1.7
assign (resid 114 and name HD*) (resid 115 and name HA) 2.0 0.2 5.9
assign (resid 114 and name HN) (resid 116 and name HN) 2.0 0.2 2.5
assign (resid 114 and name HA) (resid 117 and name HB*) 2.0 0.2 2.5
assign (resid 114 and name HD*) (resid 117 and name HD*) 2.0 0.2 5.4
assign (resid 114 and name HD*) (resid 117 and name HG*) 2.0 0.2 5
assign (resid 115 and name HG1*) (resid 115 and name HA) 2.0 0.2 1.5
assign (resid 115 and name HG2*) (resid 115 and name HA) 2.0 0.2 1.5
assign (resid 115 and name HN) (resid 115 and name HB) 2.0 0.2 0.5
assign (resid 115 and name HN) (resid 115 and name HG1*) 2.0 0.2 2.2
assign (resid 115 and name HN) (resid 115 and name HG2*) 2.0 0.2 2.2
assign (resid 115 and name HG*) (resid 116 and name HN) 2.0 0.2 3.2
assign (resid 115 and name HG1*) (resid 116 and name HN) 2.0 0.2 3.0
assign (resid 115 and name HG2*) (resid 116 and name HN) 2.0 0.2 3.0
assign (resid 115 and name HN) (resid 116 and name HN) 2.0 0.2 1.5
assign (resid 115 and name HN) (resid 117 and name HN) 2.0 0.2 2.5
assign (resid 115 and name HA) (resid 118 and name HB*) 2.0 0.2 6
assign (resid 115 and name HA) (resid 118 and name HN) 2.0 0.2 1.5
assign (resid 115 and name HG1*) (resid 119 and name HD1*) 2.0 0.2 4.0
assign (resid 115 and name HG2*) (resid 119 and name HD1*) 2.0 0.2 4.0
assign (resid 115 and name HA) (resid 119 and name HD1*) 2.0 0.2 4.5
assign (resid 115 and name HA) (resid 119 and name HN) 2.0 0.2 3.0
assign (resid 115 and name HA) (resid 119 and name HN) 2.0 0.2 3.0
assign (resid 116 and name HA) (resid 116 and name HG1) 2.0 0.2 1.7
assign (resid 116 and name HA) (resid 116 and name HG2) 2.0 0.2 1.7
assign (resid 116 and name HG*) (resid 116 and name HN) 2.0 0.2 2.5
assign (resid 116 and name HN) (resid 117 and name HN) 2.0 0.2 1.5
assign (resid 116 and name HA) (resid 119 and name HB) 2.0 0.2 1.5
assign (resid 116 and name HA) (resid 119 and name HD1*) 2.0 0.2 3.0
assign (resid 116 and name HA) (resid 119 and name HG11) 2.0 0.2 3.0
assign (resid 116 and name HA) (resid 119 and name HG12) 2.0 0.2 3.0
assign (resid 116 and name HA) (resid 119 and name HG2*) 2.0 0.2 4.2
assign (resid 116 and name HA) (resid 119 and name HN) 2.0 0.2 3.0
assign (resid 116 and name HB*) (resid 119 and name HD1*) 2.0 0.2 5.5
assign (resid 117 and name HB*) (resid 117 and name HN) 2.0 0.2 1.5
assign (resid 117 and name HG1) (resid 117 and name HN) 2.0 0.2 1.5
assign (resid 117 and name HG2) (resid 117 and name HN) 2.0 0.2 3.0
assign (resid 117 and name HB*) (resid 118 and name HN) 2.0 0.2 2.5
assign (resid 117 and name HA) (resid 120 and name HN) 2.0 0.2 1.5
assign (resid 118 and name HA) (resid 119 and name HG1*) 2.0 0.2 4.5
assign (resid 118 and name HB*) (resid 119 and name HN) 2.0 0.2 3.0
assign (resid 118 and name HN) (resid 119 and name HN) 2.0 0.2 1.5

assign (resid 118 and name HA) (resid 121 and name HN) 2.0 0.2 3.0
assign (resid 119 and name HA) (resid 119 and name HD*) 2.0 0.2 3.0
assign (resid 119 and name HA) (resid 119 and name HG2*) 2.0 0.2 2.2
assign (resid 119 and name HA) (resid 119 and name HG2*) 2.0 0.2 2.7
assign (resid 119 and name HB) (resid 119 and name HD*) 2.0 0.2 2.2
assign (resid 119 and name HG*) (resid 119 and name HA) 2.0 0.2 2.5
assign (resid 119 and name HN) (resid 119 and name HB) 2.0 0.2 0.7
assign (resid 119 and name HN) (resid 119 and name HD1*) 2.0 0.2 3.0
assign (resid 119 and name HN) (resid 119 and name HG11) 2.0 0.2 0.5
assign (resid 119 and name HN) (resid 119 and name HG12) 2.0 0.2 0.5
assign (resid 119 and name HN) (resid 119 and name HG2*) 2.0 0.2 3.0
assign (resid 119 and name HB) (resid 120 and name HN) 2.0 0.2 0.5
assign (resid 119 and name HB) (resid 120 and name HN) 2.0 0.2 1.5
assign (resid 119 and name HD1*) (resid 120 and name HN) 2.0 0.2 4.5
assign (resid 119 and name HG11) (resid 120 and name HN) 2.0 0.2 3.0
assign (resid 119 and name HG12) (resid 120 and name HN) 2.0 0.2 3.0
assign (resid 119 and name HG2*) (resid 120 and name HN) 2.0 0.2 3.0
assign (resid 119 and name HN) (resid 120 and name HN) 2.0 0.2 1.5
assign (resid 119 and name HA) (resid 122 and name HD*) 2.0 0.2 5
assign (resid 120 and name HB*) (resid 120 and name HN) 2.0 0.2 1.7
assign (resid 120 and name HN) (resid 121 and name HN) 2.0 0.2 1.5
assign (resid 120 and name HA) (resid 123 and name HB*) 2.0 0.2 4
assign (resid 120 and name HA) (resid 123 and name HG*) 2.0 0.2 4
assign (resid 121 and name HB*) (resid 121 and name HN) 2.0 0.2 2.5
assign (resid 121 and name HB*) (resid 122 and name HN) 2.0 0.2 4.0
assign (resid 121 and name HN) (resid 122 and name HN) 2.0 0.2 1.5
assign (resid 122 and name HB1) (resid 122 and name HN) 2.0 0.2 1.5
assign (resid 122 and name HB2) (resid 122 and name HN) 2.0 0.2 1.5
assign (resid 122 and name HB1) (resid 123 and name HN) 2.0 0.2 3.0
assign (resid 122 and name HB2) (resid 123 and name HN) 2.0 0.2 3.0
assign (resid 122 and name HA) (resid 124 and name HN) 2.0 0.2 1.7
assign (resid 123 and name HB1) (resid 123 and name HN) 2.0 0.2 1.7
assign (resid 123 and name HB2) (resid 123 and name HN) 2.0 0.2 1.7
assign (resid 123 and name HG*) (resid 123 and name HA) 2.0 0.2 0.5
assign (resid 123 and name HG*) (resid 123 and name HN) 2.0 0.2 4.0
assign (resid 123 and name HA) (resid 124 and name HN) 2.0 0.2 1.5