LBNL-39116
UC-408

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Structural Studies of the Activation of the Two Component Receiver Domain NTRC by Multidimensional Heteronuclear NMR

M.J. Nohaile
Structural Biology Division

May 1996
Ph.D. Thesis

RB

LBNL-39116
UC-408

Structural Studies of the Activation of the Two Component Receiver Domain NTRC by Multidimensional Heteronuclear NMR

Michael James Nohaile
Ph.D. Thesis

Department of Chemistry
University of California, Berkeley

and

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

May 1996

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

Structural Studies of the Activation of the Two Component Receiver Domain NTRC by Multidimensional Heteronuclear NMR

by

Michael James Nohaile

B.S. in Chemistry (Massachusetts Institute of Technology) 1990
B.S. in Biology (Massachusetts Institute of Technology) 1990

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA, BERKELEY

Committee in charge:

Professor David E. Wemmer, Chair
Professor Michael Botchan
Professor Raymond Stevens

1996

Structural Studies of the Activation of the Two Component Receiver Domain NTRC by Multidimensional Heteronuclear NMR

by

Michael James Nohaile

Dedicated to my father

George Abraham Nohaile

who has supported

me unfailingly in all my

endeavors, scientific and

otherwise.

Table of Contents

List of Figures

Acknowledgements

I would like to thank Dave Wemmer for providing a great environment in which to my graduate work. Dave is an excellent scientist and I am constantly amazed by the breadth and depth of his intellect.

I would like to thank Brian Volkman and Dorothee Kern for their help on the NTRC project. All of the work in Chapter 2 was done with Brian. Most of the work in chapters 3, 4, and 5 was done with the assistance of Doro. Sydney Kustu, Ken Stedman, and Nancy Amy have been excellent collaborators and provided an important biological perspective to the NTRC project.

The people of the Wemmer group have been a great source of support and friendship. Jeff Pelton, Fredd Damberger and Andrew Lee provided invaluable advice and helped form a lively protein structure group. I would like to thank the following members of the Wemmer group, past and present, for their friendship and discussion: John Cave, Ho Cho, Tim Cole, Mark Coy, Patty Fagan-Jones, Bernhard Geierstanger, Nat Gordon (who was a great help in the determination of the dynamics of NTRC in Chapter 3), Jonathan Heller, Werner Kremer, Mark Kubinec, Corey Liu, Ping Lyu, Marty Pagel, Rafa Peleaz, and Qing Zhang.

I would like to thank Stepahnie Robertson for her love and support thoughout the last two years of my graduate career. She is the most amazing person I know.

Finally, I would like to thank my family for their everlasting love and support.

Abstract

Structural Studies of the Activation of the Two Component Receiver Domain NTRC by Multidimensional Heteronuclear NMR

by

Michael James Nohaile

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor David E. Wemmer, Chair

Multidimensional heteronuclear NMR spectroscopy was used to investigate the N-terminal domain of the transcriptional enhancer NTRC (NiTrogen Regulatory protein C). This domain belongs to the family of receiver domains of two-component regulatory systems involved in signal transduction. Phosphorylation of NTRC at D54 leads to an activated form of the molecule which stimulates transcription of genes involved in nitrogen regulation. Three and four dimensional NMR techniques were used to determine an intermediate resolution structure of the unphosphorylated, inactive form of the N-terminal domain of NTRC. The structure is comprised of five α-helices and a five-stranded β-sheet in a $(\beta/\alpha)_5$ topology. Analysis of the backbone dynamics of NTRC indicate that helix 4 and strand 5 are significantly more flexible than the rest of the secondary structure of the protein and that the loops making up the active site are flexible. The short lifetime of phospho-NTRC hampers the study of this form. However, conditions for determining the resonance assignments and, possibly, the three dimensional structure of phosphorylated NTRC have been obtained. Tentative assignments of the phosphorylated form indicate that

the majority of the changes that NTRC experiences upon phosphorylation occur in helix 3, strand 4, helix 4, strand 5, and the loop between strand 5 and helix 5 (the "3445" face of NTRC) as well as near the site of phosphorylation. In order to examine a stable, activated form of the protein, constitutively active mutants of NTRC were investigated. The conformational changes in the mutants were probed by comparing the chemical shifts of the wildtype and mutant proteins. The changes seen in the mutants were generally consistent with the changes seen in the phosphorylated form. Interestingly, constitutive mutations away from the active site cause conformational changes in the active site while a constitutive mutation in the active site causes changes in the "3445" face of NTRC. This suggests that the constitutive mutations shift the protein towards the active conformation and do not just cause local changes in structure. Furthermore, this analysis indicates that the "3445" face of NTRC is important for activation.

Approved: David D. Weiss 5/13/96

Date

Chapter 1 NMR Theory and Methodology

Introduction

The determination of the three dimensional structure of a protein at atomic resolution is an important step in the understanding of the details of its function. This sort of data, along with a great deal of other experimental information, can help to explain the molecular basis of enzymatic activity, substrate specificity, quaternary associations, inhibitor binding, regulation by post-translational modifications and a host of other processes. For many years, X-ray crystallography was the sole technique capable of determining atomic resolution structures of macromolecules. However, within the last twenty years, high resolution nuclear magnetic resonance (NMR) has arisen as an alternative technique for this purpose. It is worth noting that these are in many senses complementary techniques. In contrast to X-ray crystallography, NMR is a solution technique that is capable of determining not only structural details, but also kinetic and dynamic information. On the other hand, structure determination by NMR is limited to relatively small proteins compared to those that are accessible by X-ray crystallography.

The first NMR spectrum of a protein, ribonuclease, was reported in 1957 (Saunders et al., 1957). However, it was not until the development of Fourier transform NMR techniques (Ernst & Anderson, 1966) that detailed investigation of protein structure could be done. The development of the 2 dimensional (2D) NOESY experiment (Jeener et al., 1979) and the 2D COSY experiment (Jeener, 1971; Aue et al., 1976) allowed the structures of peptides and small proteins to be determined. Quite recently, isotope labeling and the development of 3 and 4D NMR experiments (Bax, 1994) have allowed determination of proteins up to 30 kD. Other techniques, such as partial deuteration, might increase this limit up to 35-40 kD.

This chapter uses the product operator formalism and relaxation theory to describe the ^1H-^1H NOESY, COSY, and TOCSY experiments and the heteronuclear 2, 3 and 4D experiments used for structure determination. Furthermore, it describes the manner in

which the information provided by these experiments is used for the determination of protein structure.

The Assignment and Structure Problem

In order to determine the three dimensional structure of proteins via NMR, two types of information - sequence specific assignments and interproton distances - are required. Obtaining sequence specific assignments involves matching each peak in the spectrum with a particular nucleus in the protein. This gives a tag for the behavior of that particular nucleus. The distance information arises from correlations between protons that are close in space. If enough pairwise distance relationships can be determined, the overall structure of the protein can be calculated. Note that the ability to obtain distance information is dependent on having the sequence specific assignments.

Two basic types of experiments will be exploited to give the necessary information. The first makes correlations between spins that are connected through a small number of bonds. These are the COSY (Jeener, 1971; Aue et al., 1976) or TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) type experiments in the homonuclear case and the HSQC (Bodenhausen & Ruben, 1980) and HMQC (Mueller, 1979) experiments in the heteronuclear case. These types of experiments, with the exception of the TOCSY, can be explained quite well with the product operator formalism. The second type of experiment, the NOESY, gives correlations between protons that are less than 5 Å apart in space. This type of experiment is based on relaxation phenomena.

Theory and Methodology

The Basic Phenomena

Given a spin-1/2 nuclei in an applied magnetic field, B_0, there are two possible states. In one state, called the α state, the spin is aligned with the magnetic field. This is the lower energy state and will be denoted $|+\rangle$. In the other state, called the β state, the spin is aligned against the external magnetic field. This is the higher energy state and is denoted $|-\rangle$. The energy difference between the states is related to the strength of the

external magnetic field, Bo, and an characteristic constant for a given nucleus, the gyromagnetic ratio or γ:

$$\Delta E = \gamma \hbar B_o \tag{1.1}$$

where $\hbar$ is Planck's constant (6.62608×10^{-34} Js) divided by 2π. Thus, for the a B_o field of 11.7 T and a γ for protons of 2.6753×10^{8} $(Ts)^{-1}$ the energy difference between the two levels is 3.3×10^{-25} J.

If we have a large number of spin-1/2 nuclei, they will be distributed between the two energy states according to the Boltzmann equation:

$$N_-/N_+ = e^{-\Delta E/kT} \tag{1.2}$$

where N_-/N_+ is the ratio of the populations of the higher and lower energy states, k is the Boltzmann constant, 1.38×10^{-23} JK^{-1}, and T is the temperature. For protons at 298° K and an external magnetic field of 11.7 T, the ratio N_-/N_+ is 0.999919. This means that for a million spins the population difference between the states is about 40 spins. This small population difference leads to an inherent insensitivity for NMR compared to many other sorts of spectroscopy. However, the signal intensity lost due to the small ΔE is compensated by the higher resolution this engenders. According to the Heisenberg uncertainty principle, because ΔE is small, the lifetime of nonequilibrium populations is long. This extended lifetime leads to better resolved lines. This will be discussed more extensively in the section on transverse magnetization and linewidths.

The Density Matrix

Classical methods of describing NMR experiments, such as the Bloch equations, only deal with the observable magnetization. This is sufficient for some, but by no means all, experiments (Sorenson et al., 1983). In contrast, the quantum mechanical description

of NMR deals directly with the state of the system without special distinction for the final observable. There are many treatments available for the density matrix and its relation to NMR (Fano, 1957; Blum, 1981; Goldman, 1988; Munowitz, 1988; Mateescu & Valeriu, 1993; Farrar & Harriman, 1995).

The two states we have described, $|+\rangle$ and $|-\rangle$, are eigenstates of the Zeeman Hamiltonian. This Hamiltonian describes the interaction of the spins with the external magnetic field, B_o. We can write a wavefunction which describes the state of a spin in terms of the eigenstates of the Zeeman Hamiltonian. A wavefunction, $|\Psi\rangle$, can also be described as a vector corresponding to the contributions of the two eigenstates to the wavefunction:

$$|\Psi\rangle = 1*|-\rangle + 0*|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.3}$$

Thus, these eigenstates form a basis set. We can describe any wavefunction as a linear combination of this or some other basis set. In general,

$$|\Psi\rangle = c_1|\Psi_1\rangle + c_2|\Psi_2\rangle + \ldots = \sum_i c_i|\Psi_i\rangle \tag{1.4}$$

A wavefunction which is an equal mixture of the two states $|+\rangle$ and $|-\rangle$ could be generated. To do this, the two states must be coupled in such a way as to allow transitions between them. Mathematically, the states can be coupled by applying a Hamiltonian that does not commute with the Zeeman Hamiltonian. For instance, we might apply a second magnetic field, B_1, perpendicular to the external magnetic field, B_o. From the point of view of traditional spectroscopy, this can thought of as putting energy into the system at a

frequency which causes transitions between the $|+\rangle$ and $|-\rangle$ states. This frequency is, of course, determined by:

$$\Delta E = \hbar\nu \tag{1.5}$$

A mathematical construct called the density operator, ρ, can be defined to allow easy description of NMR experiments. First, the wavefunction of a single spin, Ψ, is expanded as a linear combination of some complete orthonormal basis set as described in (1.4). Next we take the outer or tensor product of the expanded wavefunction.

$$\rho = |\psi\rangle\langle\psi| = \sum_i \sum_j c_i^{*} c_j |\psi_i\rangle\langle\psi_j| \tag{1.6}$$

Note that the * symbol denotes the complex conjugate of a basis wavefunction. Now two matrices have been formed. One matrix corresponds to each member of the basis set and one matrix corresponds to the coefficients for each member of the basis set. The matrix of coefficients is called the density operator. This notation is extremely convenient because the matrix of the elements of the basis set is always the same for a given basis set. Thus, we only need to keep track of the density operator in calculations. We can reconstruct equation 1.6 at any time from that matrix and the basis set.

The pure state of a spin in an external magnetic field can be described using the density operator and the Zeeman Hamiltonian basis set. For instance, in the case where a spin is in the $|+\rangle$ state, the density operator is:

$$\rho = |\psi_\alpha\rangle\langle\psi_\alpha| = \begin{pmatrix} c_\alpha^{*} c_\alpha & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{1.7}$$

Note that the product of the coefficients must be equal to 1 since we are using a normalized basis set (this implies the spin exists). This matrix is a pure state of one spin or a state where all the spins are behaving in exactly the same manner.

In NMR, we are interested in very large ensembles of spins. This can be described by:

$$\bar{\rho} = \overline{|\psi\rangle\langle\psi|} = \sum_i \sum_j \overline{c_i^* c_j} |\psi_i\rangle\langle\psi_j| \tag{1.8}$$

This equation describes a statistical average over all possible states of the system. The matrix of coefficients for this system is called the density matrix $\underset{\approx}{\rho}$:

$$\underset{\approx}{\rho} = \overline{c_i^* c_j} \tag{1.9}$$

The diagonal elements of the density matrix describe the populations of spins in the various states of the basis set. In our case, the diagonal would correspond to the probability of find a spin in either the $|+\rangle$ or $|-\rangle$ state. The off diagonal elements correspond to the probabilities of coherent superpositions between the states.

The density matrix can be used to calculate the expectation value of particular property A for an ensemble. We start by taking the expectation values of the operator of interest in the in the basis set:

$$\overline{\langle A \rangle} = \overline{\langle \psi | \hat{A} | \psi \rangle} = \sum_i \sum_j \overline{c_i^* c_j} \langle \psi_j | \hat{A} | \psi_i \rangle \tag{1.10}$$

The density matrix naturally falls out in this analysis. The term $\langle \psi_j | \hat{A} | \psi_i \rangle$ also forms a matrix $\underset{\approx}{A}$. The expectation value of the observable A can be calculated by taking the sum of the diagonal elements or the trace of the product of $\underset{\approx}{\rho}$ and $\underset{\approx}{A}$

$$\overline{\langle A \rangle} = \mathrm{Tr}\left(\underset{\approx}{A}\,\underset{\approx}{\rho} \right) \tag{1.11}$$

This is very convenient because the matrices $\underset{\approx}{A}$ can be tabulated for various observables and basis stets. This reduces the problem to that of finding the value of $\underset{\approx}{\rho}$.

Since we have chosen to work in the Zeeman Hamiltonian basis set, we can derive the equilibrium density matrix from the Boltzmann equation. With a large number of spins, density operator at equilibrium, $\hat{\rho}_o$, is:

$$\hat{\rho}_o = \frac{1}{Z} e^{\left(-\hbar\hat{\mathcal{H}}/kT \right)} \tag{1.12}$$

where k is Boltzmann's constant, T is temperature, $\hat{\mathcal{H}}$ is the Zeeman Hamiltonian and Z is the partition function:

$$Z = \mathrm{Tr}\left(e^{\left(-\hbar\hat{\mathcal{H}}/kT \right)} \right) \tag{1.13}$$

Note that $\hat{\rho}_o$ is the operator form of the density matrix and, as such, $\hat{\mathcal{H}}$ represents $E/\hbar$. The density operator can be written as a Taylor expansion. Since the energies involved in NMR are very small compared to kT, we can truncate this form to yield the high temperature approximation:

$$\hat{\rho}_o \approx \frac{1}{Z}\left(\hat{1} - \hbar/kT\,\hat{\mathcal{H}}\right) \tag{1.14}$$

with the partition function:

$$Z \approx \mathrm{Tr}\left(\hat{1}\right) = N \tag{1.15}$$

where $\hat{1}$ is the identity operator and N is the dimension of the space spanned by the basis set. Thus, the density operator at equilibrium is simply proportional to the Zeeman Hamiltonian.

Next, we need to calculate how the density operator can change under the influence of another Hamiltonian. The differential equation of motion for the density operator, also known as the Liouville - von Neumann equation, is:

$$\frac{d}{dt}\hat{\rho}(t) = -i\left[\hat{\mathcal{H}}, \hat{\rho}(t)\right] \tag{1.16}$$

Thus, in order for the system to evolve, the applied Hamiltonian and the density operator must not commute. The solution to the Liouville - von Neumann equation is:

$$\hat{\rho}(t) = e^{-i\hat{\mathcal{H}}t}\hat{\rho}(0)e^{+i\hat{\mathcal{H}}t} \tag{1.17}$$

This equation states that if the density operator at time 0 (equilibrium) and the applied Hamiltonian are known, the density operator can be calculated at any time, t, later. Since $\hat{\rho}(0)$ was calculated in (1.14), all we need to describe an NMR experiment are the relevant Hamiltonians. These will be discussed shortly.

First, the question of how to apply an operator (the Hamiltonian) which appears as the argument of an exponential, (1.17), must be discussed. In order to answer this

question, the density operator will be reformulated in terms of the angular momentum operators. The spin of a nucleus gives rise to an internal angular momentum. There are three independent operators of spin angular momentum in the Cartesian frame: $\hat{I}_z$, $\hat{I}_y$, and $\hat{I}_x$. These operators act on wavefunctions in our the basis set as follows:

$$\begin{aligned} \hat{I}_x|+\rangle &= \hbar/2|-\rangle \quad \hat{I}_y|+\rangle = i\hbar/2|-\rangle \quad \hat{I}_z|+\rangle = \hbar/2|+\rangle \\ \hat{I}_x|-\rangle &= \hbar/2|+\rangle \quad \hat{I}_y|-\rangle = -i\hbar/2|+\rangle \quad \hat{I}_z|-\rangle = -\hbar/2|-\rangle \end{aligned} \tag{1.18}$$

The spin operators can be written as matrices:

$$\hat{I}_z = \hbar/2\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \hat{I}_y = i\hbar/2\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \hat{I}_x = \hbar/2\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{1.19}$$

The commutation relations of the spin operators are cyclic permutable:

$$\left[\hat{I}_z, \hat{I}_x\right] = i\hat{I}_y \qquad \left[\hat{I}_z, \hat{I}_y\right] = i\hat{I}_x \qquad \left[\hat{I}_x, \hat{I}_y\right] = i\hat{I}_z \tag{1.20}$$

Finally, we need the squares of the spin operators which are:

$$\hat{I}_x{}^2 = \frac{\hbar^2}{4}\hat{1} \qquad \hat{I}_y{}^2 = \frac{\hbar^2}{4}\hat{1} \qquad \hat{I}_z{}^2 = \frac{\hbar^2}{4}\hat{1} \tag{1.21}$$

where $\hat{1}$ is the unity operator.

Now, if we apply a Hamiltonian that is written in terms of the spin operators, we can utilize (1.17). For instance, the Hamiltonian for a radiofrequency pulse is $\Theta\hat{I}_x$,where Θ is $\gamma B_1 \tau$. If we put this Hamiltonian into one of the exponentials in (1.17) and do an expansion, we get:

$$e^{\Theta\hat{I}_x} = \hat{1} + i\Theta\hat{I}_x - \frac{\Theta^2}{2!}\hat{I}_x{}^2 - \frac{\Theta^3}{3!}\hat{I}_x{}^3 + \ldots \qquad (1.22)$$

By grouping together the even and odd powers of $\hat{I}_x$, using the values for the squares of the spins operators (1.21), and using the properties of the Taylor expansion, we get:

$$e^{\Theta\hat{I}_x} = \hat{1}\cos\Theta/2 + 2i\hat{I}_x \sin\Theta/2 \qquad (1.23)$$

When the entirety of (1.17) is evaluated it leads to:

$$\rho(t) = \left[\hat{1}\cos\Theta/2 - 2i\hat{I}_x \sin\Theta/2\right]\rho(0)\left[\hat{1}\cos\Theta/2 + 2i\hat{I}_x \sin\Theta/2\right] \qquad (1.24)$$

This equation can be easily evaluated in matrix form for particular values of Θ. What is particularly nice about this formulation is that the manipulations all appear as rotation operations. For instance, the radiofrequency pulse applied along the x-axis in (1.24) yields a rotation about that axis away from the z-axis towards the y-axis.

The Product Operator Formalism

The product operator formalism is an extremely convenient representation of the density operator for NMR (Packer & Wright, 1983; Sorenson et al., 1983; Van De Ven & Hilbers, 1983; Howarth et al., 1986). In this formalism, we will use the Cartesian spin operators, $\hat{I}_x$, $\hat{I}_y$, and $\hat{I}_z$, plus the identity matrix, $\hat{1}$, to represent the elements of the density matrix. In general, the density operator for a single spin 1/2 nuclei can be described by linear combinations of these spin operators. What makes this particularly attractive is that operations by these operators on one another correspond to rotations as in the case described in (1.23). Thus, this formalism retains some the intuitive appeal of simple vector based approaches. For instance, the density matrix at equilibrium

corresponds to $\hat{I}_z$ which is in line with a vector picture of bulk magnetization along the z-axis at equilibrium.

The discussion above is sufficient for a single spin 1/2 nuclei. However, for a pair of weakly coupled spins, the Cartesian spin operators are an insufficient basis set to describe the density operator. For this case, we need to add the products of the Cartesian spin operators: $\hat{1}$, $\hat{I}_x$, $\hat{I}_y$, $\hat{I}_z$, $\hat{S}_x$, $\hat{S}_y$, $\hat{S}_z$, $2\hat{I}_z\hat{S}_z$, $2\hat{I}_y\hat{S}_y$, $2\hat{I}_x\hat{S}_x$, $2\hat{I}_z\hat{S}_x$, $2\hat{I}_z\hat{S}_y$, $2\hat{I}_x\hat{S}_y$, $2\hat{I}_y\hat{S}_x$, $2\hat{I}_x\hat{S}_z$, and $2\hat{I}_y\hat{S}_z$. The notation $\hat{S}$ indicates the other spin in the weakly coupled system. Note that these operators are constructed by taking the outer product or tensor product of the single spin operators.

The operators of the form $\hat{I}_z$ correspond to longitudinal magnetization (aligned with the external field, B_o, along the z-axis). Operators of the form $\hat{I}_x$ or $\hat{I}_y$ correspond to in phase transverse magnetization. An operator the form $2\hat{I}_y\hat{S}_z$ corresponds to antiphase magnetization of spin I with respect to spin S. An operator of the form $2\hat{I}_z\hat{S}_z$ corresponds to longitudinal two spin order of spins S and I. Finally, an operator of the form $2\hat{I}_x\hat{S}_y$ corresponds to a two spin coherence of spins I and S.

The Hamiltonians of NMR

In order to make use of (1.17) to describe NMR experiments, the Hamiltonian must be known. In high resolution liquid NMR there are three Hamiltonian that are encountered. The first is the Hamiltonian of a radiofrequency pulse. This can be thought of as a rotating magnetic field, B_1, applied perpendicularly to the external magnetic field B_o. This Hamiltonian has the form:

$$\hat{\mathcal{H}} = (\gamma B_1)\hat{I}_n \qquad (1.25)$$

where n is the axis along which the pulse is applied. This Hamiltonian will cause the magnetization to rotate about this axis n.

The second Hamiltonian describes the chemical shift. This effect arises from small variations in the local magnetic environment which can modify the strength of the external magnetic field at a particular spin:

$$\hat{\mathcal{H}} = (\Omega_I)\hat{I}_z \tag{1.26}$$

where Ω_I is the offset of spin I form the carrier frequency. This Hamiltonian can be thought of as causing a rotation about the z-axis.

The third is the scalar or J coupling Hamiltonian. This phenomena is caused by the interaction of two spins mediated by bonding electrons. The scalar coupling Hamiltonian between two spins I and S is:

$$\hat{\mathcal{H}} = 2\pi J_{IS}\hat{I}_z\hat{S}_z \tag{1.27}$$

where J_{IS} is the scalar coupling constant between the two spins. This can be thought of as a rotation about the zz-axis. Of course, this only makes sense in the sixteen dimensional basis space of the product operators.

1 Dimensional NMR

With the product operator formalism developed above, NMR experiments can be discussed. The simplest experiment in FT NMR is to apply a short radiofrequency pulse to a sample at equilibrium and immediately turn on the receiver to collect the free induction decay (FID). In this case, the chemical shift Hamiltonian is active during acquisition. Using the product operator formalism, we have:

$$I_z \xrightarrow{\pi/2 I_x} I_z\cos(\pi/2) - I_y\sin(\pi/2) = -I_y \xrightarrow{(\Omega t)\hat{I}_Z} -I_y\cos(\Omega t) + I_x\sin(\Omega t) \tag{1.28}$$

Thus, we will detect the spins at their chemical shifts. Such a 1D ^1H spectrum is shown in Figure 2.6. As mentioned above, correlations between spins are required to obtain the information necessary for structure determination. This is impossible to do with 1D spectra on a protein the size of NTRC (~14 kD). Thus, 2D NMR is required.

2 Dimensional NMR

The general scheme in 2D NMR is to prepare magnetization on a starting nucleus, spin a, and label that magnetization during an incremented evolution period, t_1. This is followed by some combination of radiofrequency pulses and delays, known as the mixing time, which transfers some of the magnetization from spin a to another nucleus, spin b. Then, magnetization is detected during the acquisition period t_2. If we run this experiment with a large number of different t_1 timepoints, the magnetization that is transferred to spin b from spin a is modulated in either amplitude or phase by the frequency of spin a. After Fourier transformation, this magnetization will appear with the frequency of spin a in one dimension and the frequency of spin b in the other dimension. Thus, a correlation between spin a and spin b has been created. This is shown in Figure 1.1.

The COSY Experiment

In the case of the COSY (COrrelation SpectroscopY) experiment (Jeener, 1971; Aue et al., 1976) the correlations are based on J couplings and, as such, are through bond correlations. The pulse train for this experiment is shown in Figure 1.2A. The experimental scheme is as follows. A 90° pulse prepares the magnetization:

$$^aI_z \xrightarrow{\pi/2 I_y} {}^aI_x \qquad (1.29)$$

This is followed by the incremented t_1 evolution period. In this period both the chemical shift and J coupling Hamiltonians are active. In the COSY experiment, this t_1 period

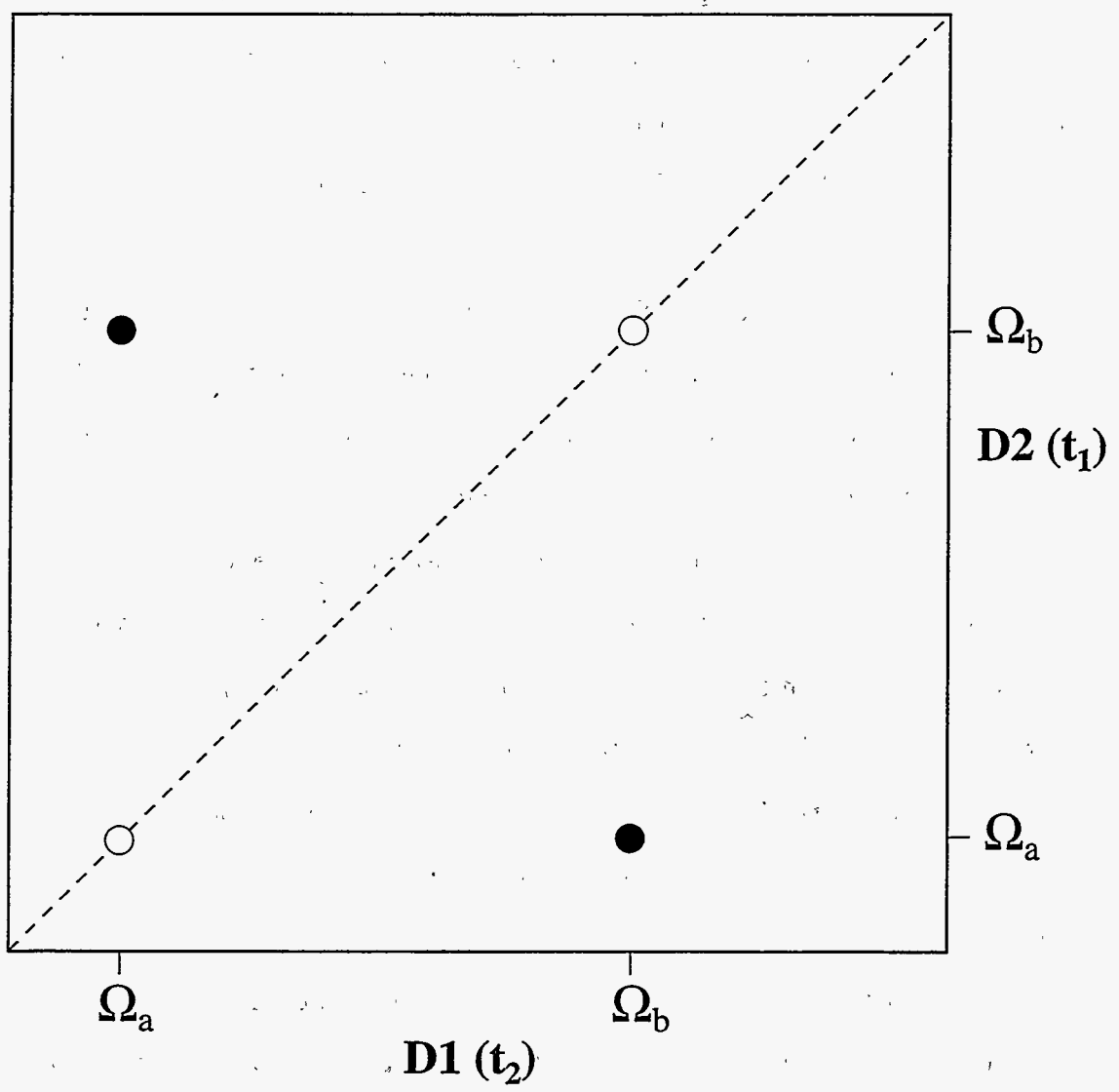


Figure 1.1: Schematic diagram of a 2D homonuclear correlation experiment. The open circeles on the diagonal correspond to peaks that are detected at the same frequency in both dimensions. These peaks arise from magnetization that does not transfer from spin a to spin b (or vice-versa) during the mixing period. The filled circles correspond to crosspeaks that have the frequency of one spin in the first dimension and the frequency of the other spin in the second dimension. These crosspeaks arise from magnetization that starts on spin a and transfers to spin b (or vice versa) during the mixing period.

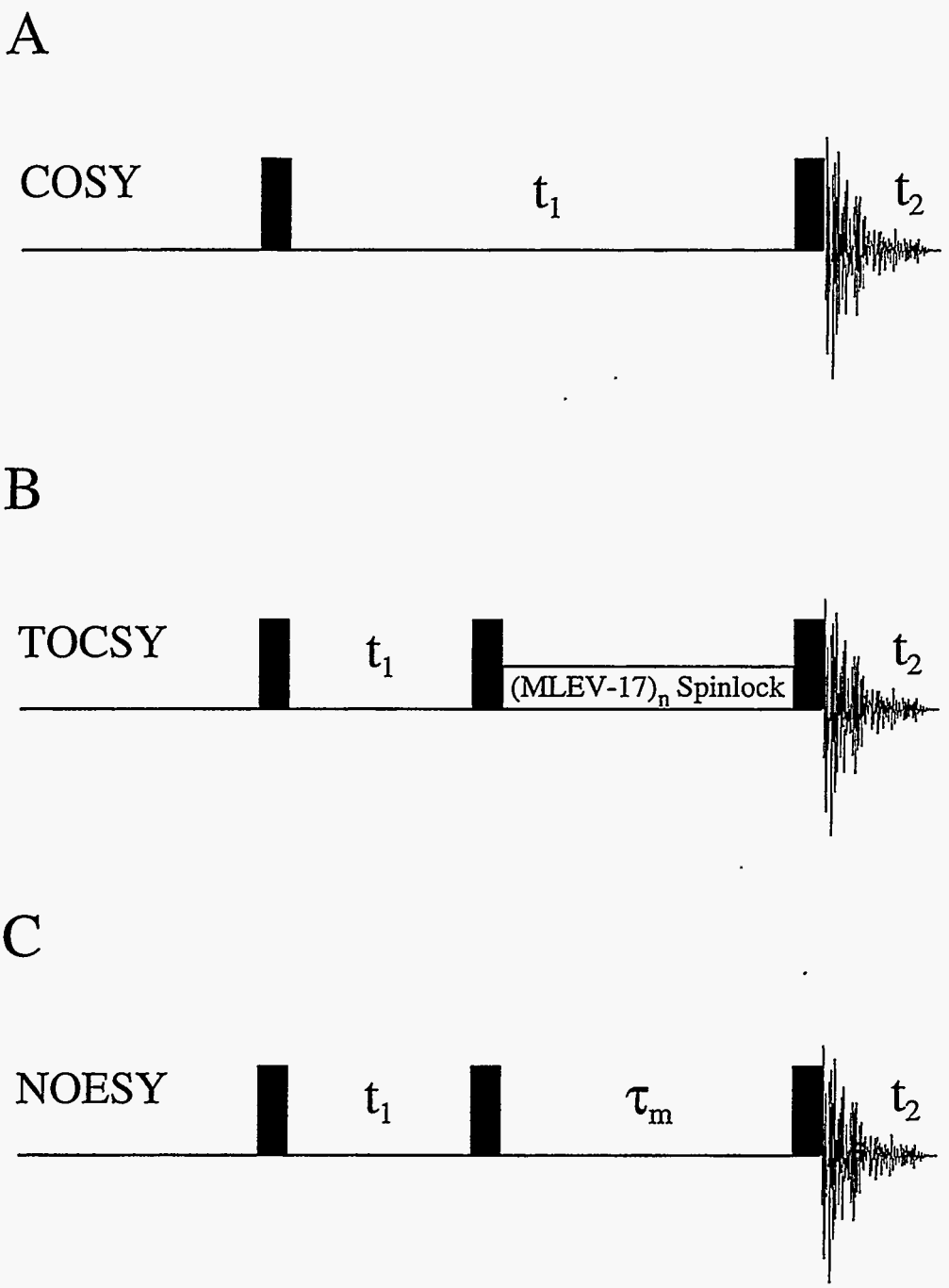


Figure 1.2: Schematic diagrams of 2D ^{1}H-^{1}H pulse sequences. $\pi/2$ pulses are indicated with filled blocks. τ_m is the the mixing time. See the text for details about these experiments.

serves the dual purpose of encoding the frequency of spin a and preparing for the mixing period by creating antiphase magnetization of spin a with respect to spin b. We now apply the chemical shift Hamiltonian:

$$ {}^{a}I_x \xrightarrow{(\Omega t)\hat{I}_z} {}^{a}I_x \cos(\Omega_a t) + {}^{a}I_y \sin(\Omega_a t) \qquad (1.30) $$

For the sake of simplicity, only the cosine term of (1.30) will be followed further. The J-coupling Hamiltonian is now applied:

$$ {}^{a}I_x \cos(\Omega_a t) \xrightarrow{J_{ab} t_1 {}^{a}\hat{I}_z {}^{b}\hat{I}_z} {}^{a}I_x \cos(\Omega_a t)\cos\left(\frac{J_{ab}}{2} t_1\right) $$
$$ + 2\,{}^{a}I_y\,{}^{b}I_z \cos(\Omega_a t_1)\sin\left(\frac{J_{ab}}{2} t_1\right) \qquad (1.31) $$

A correlation between spin a and spin b via antiphase magnetization, $2\,{}^{a}I_y\,{}^{b}I_z$, has been generated. The second pulse performs the neat trick of transferring polarization from spin a to spin b.

$$ \xrightarrow{-\pi/2 I_x} {}^{a}I_x \cos(\Omega_a t)\cos\left(\frac{J_{ab}}{2} t_1\right) $$
$$ + 2\,{}^{a}I_z\,{}^{b}I_y \cos(\Omega_a t_1)\sin\left(\frac{J_{ab}}{2} t_1\right) \qquad (1.32) $$

Thus, we have transformed magnetization on spin a, which is antiphase with respect to spin b, to magnetization on spin b, which is antiphase with respect to spin a. This is the essential step in the COSY. Now during the acquisition time, t_2, the second antiphase term in (1.32) evolves to observable magnetization under the J-coupling Hamiltonian. Note that

to get the final observed spectrum, both the J coupling and chemical shift Hamiltonians must be applied during acquisition.

The COSY experiment can be analyzed in terms of the schematic figure diagram of a homonuclear correlation experiment shown in Figure 1.1. The diagonal peaks correspond to magnetization that was labeled on one spin in t_1 and was not transferred to the other spin in t_2. The crosspeaks correspond to the situation described above where magnetization starts on spin a and is transferred to spin b via scalar coupling.

The utility of this experiment lies in its ability to correlate protons that are connected through 2, 3, or, rarely, 4 bonds. Thus, in the case of proteins, we can map out the structure of a particular amino acid by moving from proton to proton. Different amino acids will show different types of connectivities. Unfortunately, many amino acid spin system types, such as glutamic acid and glutamine, appear very similar in this experiment (Wüthrich, 1986). However, there are enough unique spin systems, such as glycine, that one can begin to get a handle on amino acid types form this experiment.

The intensity of a COSY crosspeak is determined by the balance of the buildup of transferred magnetization due to the J coupling and loss of magnetization due to relaxation (relaxation is discussed below). In general, for a protein of reasonable size (10-20 kD), the J couplings are only large enough to detect correlations between protons separated by 2 or 3 bonds. Thus, the coherences between protons on different amino acids in a protein are not seen. We will turn to the NOESY experiment to make these correlations.

The 2 Dimensional TOCSY Experiment

The drawback of the COSY experiment is that it can be quite difficult to walk through the correlations for a long sidechain. The TOCSY experiment (TOtal Correlation SpectroscopY) give the same information as the COSY experiment in a more convenient form (Braunschweiler & Ernst, 1983; Bax & Davis, 1985). In this experiment, magnetization starts on spin a and transfers to spin b, 2 or 3 bonds away, via J coupling.

However, unlike the COSY, the process continues and the transferred magnetization immediately begins to build up on spin c which is 2 or 3 bonds from spin b.

The pulse train for the TOCSY is shown in Figure 1.2B. This pulse sequence is similar to the COSY except that there is a longer TOCSY mixing period. This mixing period is generated by spin locking the magnetization in the xy plane along one axis. Alternatively, the magnetization can be locked along the z-axis to give the exact same effect (this is done in the pulse sequence shown in Figure 1.2B). This spin lock causes the spins to experience strong coupling. This means that the chemical shift difference between the two spins is smaller than the J coupling between them. This mixing period causes a transfer of magnetization between spins by cross-polarization or isotropic mixing. Unfortunately, the product operator formalism discussed above can't describe this situation adequately because it was developed for the weak coupling limit. The crux of the matter is that the simple product operator basis set is no longer convenient because of the mixing of states in the strong coupling limit. Thus, a new basis set must be introduced to describe this experiments. This basis set contains the sum and differences of the simple product operators. Furthermore, the reduced J coupling Hamiltonian in (1.27) is no longer sufficient. Instead the full J coupling Hamiltonian:

$$\hat{\mathcal{H}} = J_{IS}\left(\hat{I}_z\hat{S}_z + \hat{I}_x\hat{S}_x + \hat{I}_y\hat{S}_y\right) \tag{1.33}$$

must be used. The details of this treatment can be found elsewhere (Hicks et al., 1994). Note that, once again, there is a competition between the J-coupling and relaxation in the buildup of TOCSY crosspeaks. Therefore, spins distant from the initial spin will show less intense crosspeaks than those close to the initial spin. This can be quite problematic in large proteins with accelerated relaxation.

The ^{1}H-^{1}H 2D TOCSY experiment is extremely convenient to use because, in principle, the entire amino acid spin system is correlated to every proton in spin system.

Thus, in an aspartic acid, the complete spin system (NH, $H\alpha$, and $H\beta$'s) is correlated at the chemical shift of the NH, the $H\alpha$, and the $H\beta$'s. This greatly facilitates the determination of the spin system type.

Relaxation and the NOE

In order to describe the NOESY experiment (Jeener et al., 1979) and to lay the foundation for the dynamics data in chapter 3, this section will deal with some aspects of relaxation in NMR (Abragam, 1961; McConnell, 1987; Goldman, 1988; van de Ven, 1995).

Relaxation processes in NMR are commonly divided into two types. T_1 relaxation, also called longitudinal or spin-lattice relaxation, refers to the time it takes a sample to reach the thermal equilibrium defined in (1.2) via transference of energy to other degrees of freedom such as molecular motion. This process requires an exchange of energy with the surroundings (the lattice). This exchange of energy must be done at the discrete frequency of the energy difference between $|+\rangle$ and $|-\rangle$ states as defined by Planck's equation:

$$\Delta E = E_\beta - E_\alpha = h\nu \tag{1.33}$$

T_2 relaxation, also called transverse or spin-spin relaxation, refers to the time it takes the magnetization to lose coherence in the transverse plane. This can be thought of as a loss of the phase coherence of the off-diagonal terms in the density matrix. It is a consequence of small local perturbations in the magnetic field in the z direction. Note that T_2 relaxation can also occur due to transitions between the $|+\rangle$ and $|-\rangle$ states.

The interactions that give rise to the frequencies which cause relaxation come from anisotropic interactions of the spin with the external applied magnetic field, B_o, and with other spins. Since the molecules are tumbling randomly, these anisotropic interactions randomly fluctuate in time. This can be seen in a functional form, H(t), in Figure 1.3A. A

measure of the strength of these randomly fluctuating fields can be described in a autocorrelation function:

$$G(\tau) = \overline{H(t)H(t+\tau)} \tag{1.34}$$

which is the mean square average of the random function , H(t). This function gives a measure of self-similarity of an ensemble at a time, τ, later than the initial state. Thus, the autocorrelation function measures how long it takes a system to become uncorrelated to its previous position. As τ increases, the correlation between the ensemble at time t and time t+τ later drops exponentially. Thus, G(τ) can be modeled as an exponential decay with a time constant τ_c which is called the correlation time (Figure 1.4B). The correlation time is a measure of the rate at which the molecule is tumbling.

The strength of these anisotropic fluctuations at any particular frequency is given by the spectral density function which is the Fourier transform of the autocorrelation function.

$$J(\omega) = \int_{-\infty}^{\infty} G(\tau) e^{i\omega t} d\tau \tag{1.35}$$

This has the form of a Lorentzian:

$$J(\omega) = \frac{\tau_c}{1+\omega^2\tau_c^2} \tag{1.36}$$

The spectral density function is depicted at two different correlation times in Figure 1.4C. Note that the spectral density function is centered at 0 because the autocorrelation function does not oscillate. Also, the area under the spectral density curve is constant. Thus, as the

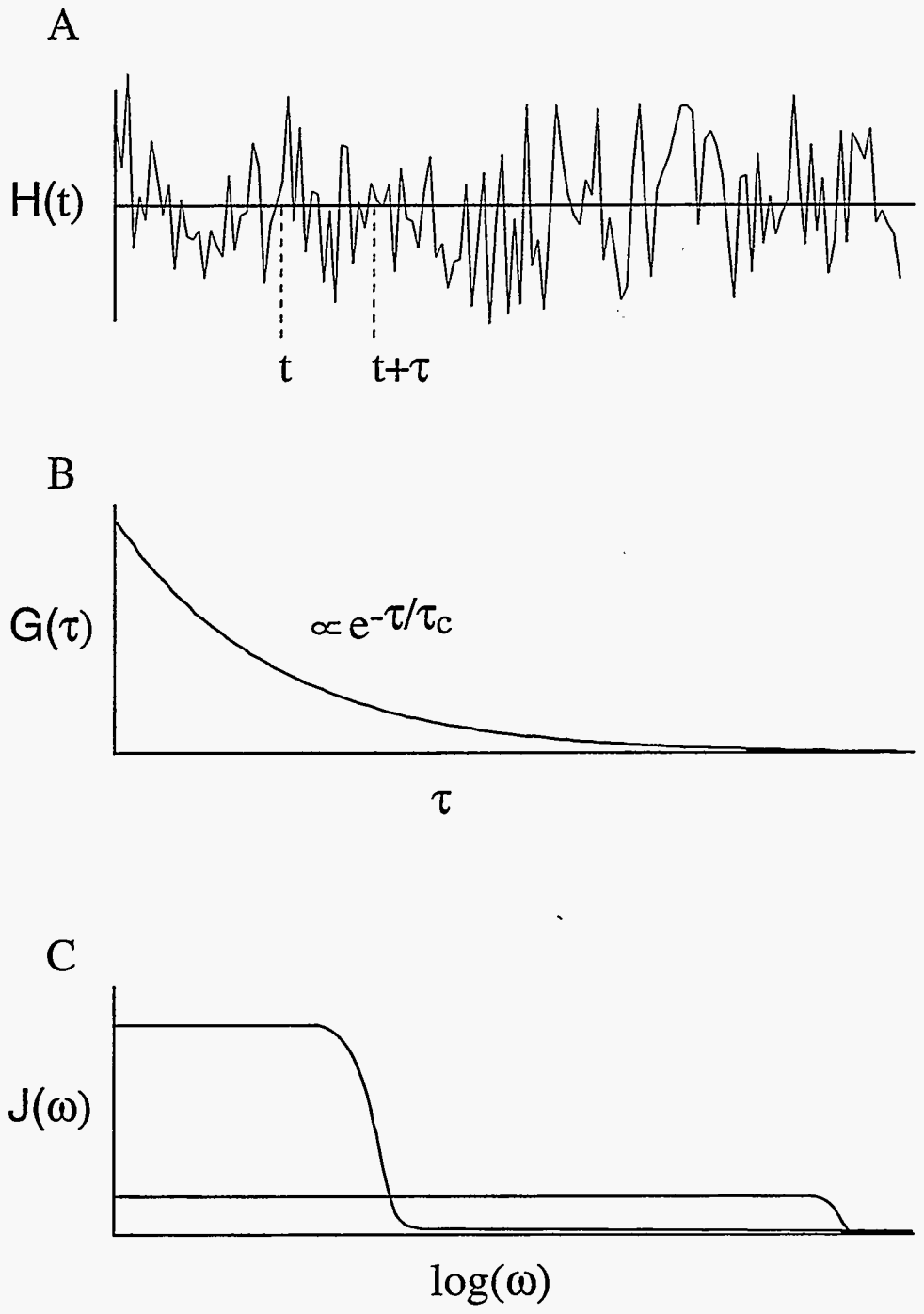


Figure 1.3: (A) A function randomly fluctuating in time. This is a model of the anisotropic interactions of a molecule rapidly tumbling in a liquid. (B) The autocorrelation function dervided from a randomly fluctuating function as a function of τ. This is a exponential decay with the correlation time, τ_c, as characteristic time constant. (C) The spectral density function for two different correlation times.

correlation time lengthens, fewer frequencies are sampled, but the intensity at each sampled frequency increases.

The functional form of T_1 relaxation in terms of the spectral density function is proportional to $J(\omega_0)$ and $J(2\omega_0)$. This is indicative of the fact that T_1 is dependent on frequencies (ω_0) which can cause transitions between the $|+\rangle$ and $|-\rangle$ states. The functional form of T_2 relaxation in terms of the spectral density function is proportional to $J(\omega_0)$ and $J(0)$. This indicates that coherence can be lost in the xy plane by transitions between the states (hence the $J(\omega_0)$ term). However, this coherence can also be lost due to small fluctuations in the magnetization field along the z-axis. These low frequency perturbations are probed by $J(0)$.

Figure 1.4 shows the dependence of T_1 and T_2 relaxation on the correlation time. Since T_1 is dependent on $J(\omega_0)$, as the correlation time grows, the spectral density at $J(\omega_0)$ increases to maximum point ($\tau_c = 1/\omega_0$) which shortens T_1. Once the correlation time passes through this point, the spectral density at $J(\omega_0)$ decreases and T_1 increases. In contrast, T_2 is mostly dependent on $J(0)$. Therefore, T_2 grows shorter as the correlation time increases. Since the correlation time increases as the molecular weight of a molecule increases, T_2 relaxation is the major barrier to NMR of large proteins. Simply put, the size of the protein under investigation increases, T_2 relaxation becomes so fast that coherences die away as they are being manipulated in pulse sequences.

The type of relaxation that we are most interested in is the dipole-dipole interaction. This phenomenon arises from the magnetic interaction of two nuclei. If we consider two spins that are coupled by the dipole-dipole interaction, we can construct the diagram shown in Figure 1.5. In this figure, a W with the appropriate subscripts (0 is zero quantum, 1 is single quantum, and 2 is double quantum) and superscripts (a indicates spin a and b indicate spin b) describe the rate constants for the various transitions. Under the selection rules of the interaction of the spins and the radiofrequency pulse the W_2 and W_0 transitions

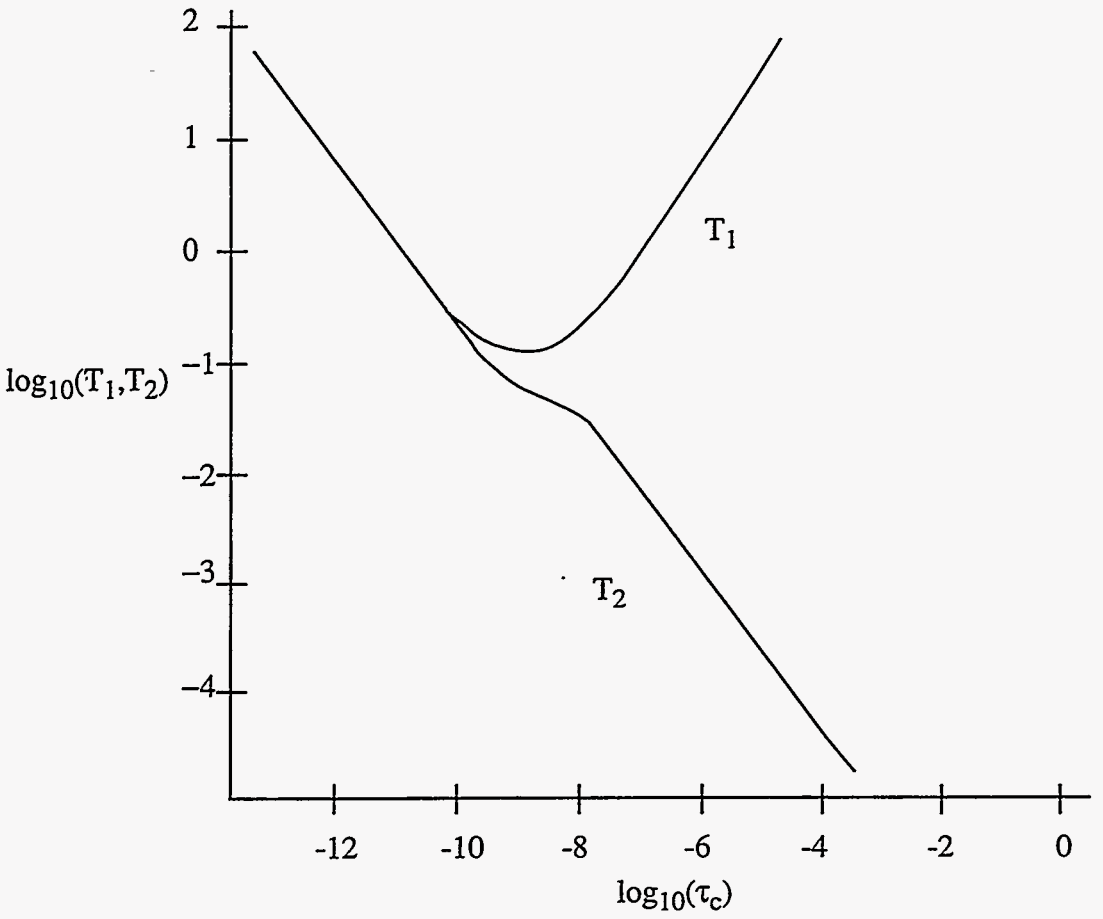


Figure 1.4: A plot of T_1 and T_2 as a function of the correlation time, τ_c, at 400 MHz.

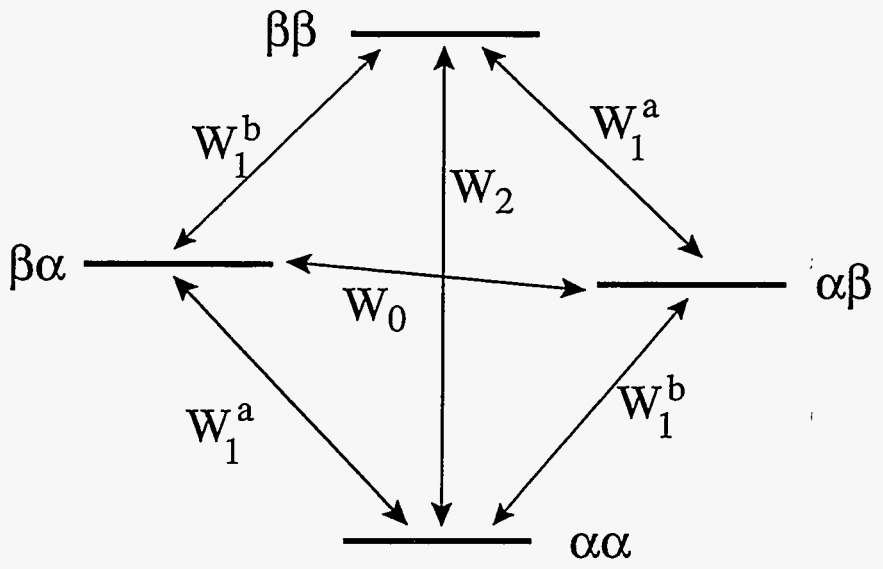


Figure 1.5: Schematic energy level diagram of a two spin-1/2 system. The rates of transition between the levels are indicated by a W with the appropriate subscript and superscript.

are forbidden. However, the selection rules allow these transitions in the case of dipolar interactions.

The dipole-dipole Hamiltonian for two identical spin-1/2 nuclei is:

$$\hat{\mathcal{H}} = \gamma_a \gamma_b \hbar^2 \left(\frac{\hat{I}_a \bullet \hat{I}_b}{r^3} - 3\frac{\left(\hat{I}_a \bullet r\right)\left(\hat{I}_b \bullet r\right)}{r^5} \right) \left(\frac{\mu_0}{4\pi} \right) \qquad (1.37)$$

where $\hat{I}$ is the nuclear spin operator, γ is the gyromagnetic ratio, r is the distance between the spins and μ_0 the permeability of a vacuum. This can be expanded into various terms that correspond to particular types of transitions (zero quantum, single quantum etc.). Note that this Hamiltonian averages exactly to zero in an isotropic solution. Therefore, it is not a Hamiltonian that must be considered in the product operator formulation for pulse sequences.

This dipolar interaction is the basis for the NOE (Nuclear Overhauser Effect) (Neuhaus & Williamson, 1989). This effect allows the transfer of polarization from one spin to another spin through space. If spin a in the coupled system shown in Figure 1.5 is saturated by a long weak radiofrequency pulse, the populations of energy levels $\alpha\alpha$ and $\beta\alpha$ and the energy levels $\alpha\beta$ and $\beta\beta$ will be equalized. This leads to a nonequilibrium situation. The system can return to equilibrium through W_0 (zero quantum) or W_2 (double quantum) transitions, as well as W_1 (single quantum).

If the W_0 pathway is the dominant mode of relaxation, the population in the $\alpha\beta$ state will increase. This cause the population difference between the α and β states to be reduced for spin b. Thus, the saturation of spin a will cause a decrease in the intensity of spin b through a dipolar coupling relaxation mechanism. Restating this, there is a negative NOE on spin b due to spin a.

If the W_2 pathway is the dominant mode of relaxation, the population in the $\alpha\alpha$ state will increase. This cause the population difference between the α and β states to be

increased for spin b. Thus, the saturation of spin a will cause an increase in the intensity of spin b through a dipolar coupling relaxation mechanism. Restating this, there is a positive NOE on spin b due to spin a.

The balance between W_0 and W_2 will determine the sign of the observed NOE. This, in turn, is determined by the balance of the spectral densities for the frequencies associated with W_0 and W_2. W_0 is dependent on low frequencies near J(0) while W_2 is dependent on large double quantum frequencies. In a large molecule with a long τ_c, the low frequencies have more spectral density than the high frequencies and W_0 dominates which gives a negative NOE. In a small molecule with a very short τ_c, the higher frequencies have a significant amount of spectral density and W_2 dominates leading to a positive NOE. For some intermediate τ_c, W_0 and W_2 can cancel out and no NOE will be observed. Proteins have very long correlation times and thus show negative NOEs.

The distance dependence for the NOE arises out of the distance dependence term in the dipolar Hamiltonian. The NOE effect falls off as $1/r^6$. In practice, relaxation due to dipolar coupling between spins that are less than 5Å apart is detectable.

The 2 Dimensional NOESY Experiment

The 2D NOESY experiment is the through space analog to the 2D COSY experiment. This experiment gives rise to crosspeaks that indicate that two spins are less than 5Å apart. An example of a 2D NOESY for NTRCis found in Figure 2.7.

The pulse train for this experiment is shown in Figure 1.2C. After an initial preparatory pulse, the magnetization is labeled on spin a during the time period t_1. A second $\pi/2$ pulse converts the magnetization back along the z-axis. Now, transfer of magnetization from spin a to spin b is caused by the NOE phenomena. Note that the magnetization is transferred back along the z-axis after frequency labeling in t_1 because, unlike the other phenomena that have been discussed, the NOE is dependent only on population differences. The third pulse puts the magnetization back in the xy plane where it

is detected. A crosspeak will arise between spin a and spin b due to the magnetization transferred from spin a to spin b during the NOESY mixing time.

The Homonuclear Assignment Problem

The first task in the NMR analysis of a protein is to determine the sequence specific assignments for all the protons. Classically, this is done with a combination of through bond experiments (COSY and TOCSY), through space experiments (NOESY) and prior knowledge of the protein sequence. This process is shown in Figure 1.6A. First, determination of spin system types is made using the COSY or TOCSY. Then the spin systems are connected through use of the NOESY. There are a number of characteristic NOE's which arise sequentially in a protein. Those most commonly used for sequential assignment are alpha proton ($H\alpha$) to amide proton (NH), NH to NH, and beta proton ($H\beta$) to NH connectives (Figure 1.6A). However, this procedure can be difficult because one cannot be certain that a particular NOE indicates a sequential connectivity and not a long range connectivity. Finally, the fragments of the protein that have been linked together in this manner are compared to the protein sequence to determine their positioning. For instance, if one has found spin systems and sequential connectivities consistent with the fragment Gly-Ala-Thr, the protein sequence can be searched for this particular combination. If the combination is unique in the sequence, then the spin systems can be assigned to that particular stretch of the protein sequence. This procedure is repeated until the entire sequence is accounted for.

Secondary structure information can be gleaned form this method. There are patterns of NOE's which are indicative of α helix and β sheet. In α helices, the following connectivities are seen: d_{NN}, $d_{\alpha N}(i, i+3)$, $d_{\alpha N}(i, i+4)$, and $d_{\alpha\beta}(i, i+3)$. In β sheet, the $d_{\alpha N}$ connectivities predominate and cross strand NOE's are seen.

This methodology is only viable for proteins under about 10 kD in weight. With proteins larger than this size, overlap in the spectrum often becomes to severe to allow analysis.

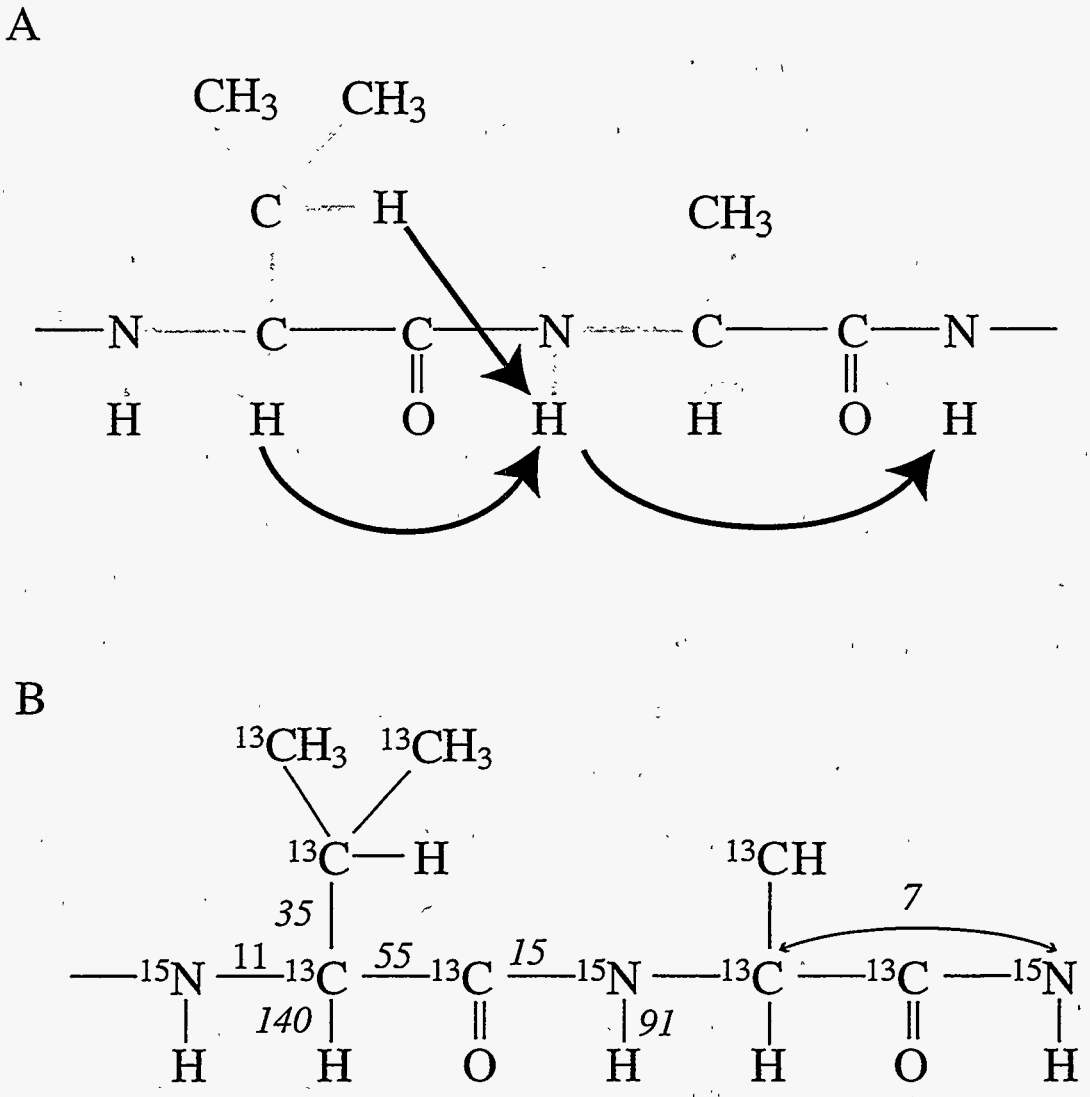


Figure 1.6: Schematic diagram of a dipeptide fragment. (A) Traditional assignment methodology. The grey bonds delineate individual spin systems which are linked in a COSY or TOCSY experiment. The arrows indicate various through space correlations seen in the NOESY which are used to link spin systems together (d_{NN},$d_{\alpha N}$, and $d_{\beta N}$). (B) Diagram of a dipeptide fragment with the large heteronuclear 1 and 2 bond couplings (in Hz) used for triple resonance experiments indicated.

.Isotopic Labeling

Recently, a revolution has taken place in protein NMR (Bax & Grzesiek, 1993). The ability to isotopically enrich proteins with low natural abundance spin 1/2 nuclei such as ^{13}C and ^{15}N has allowed new classes of experiments to be developed to investigate much larger proteins (up to 30 kD). These experiments take advantage of the spin-1/2 heteronuclei in three ways. First, these experiments use relatively large one and two bond scalar couplings between heteronuclei. These large J couplings better tolerate the shorter T_2's of large proteins. Second, the overlap of the ^1H-^1H spectra can be separated into higher dimensions by the heteronuclear chemical shifts. Finally, ^{13}C and ^{15}N have a larger chemical shift range than ^1H.

Due to the low natural abundance of ^{13}C and ^{15}N, proteins enriched with these nuclei must be obtained from recombinant sources. This can be accomplished by growing the strain of bacteria overexpressing the protein of interest in a minimal media with defined carbon, usually glucose, and nitrogen, usually ammonium chloride, sources. 98% ^{15}N ammonium chloride and 99% ^{13}C labeled glucose are commercially available. Growth on this media will typically yield samples that are nearly uniformly labeled with either or both ^{15}N and ^{13}C. With a strong promoter and a soluble protein, yields of 10-80 mg of purified protein per liter are obtainable.

It is also possible to selectively label particular amino acids (McIntosh & Dahlquist, 1990). Amino acids labeled at a particular position, such as the amide nitrogen or the alpha carbon, are commercially available and can be incorporated in a similar manner to uniform labeling. The only difference is that the selectively labeled amino acids as well as unlabeled amino acids must be added to the minimal media. The unlabeled amino acids are intended to suppress dilution of the label by through various metabolic pathways. Whether a particular amino acid is an appropriate target for selective labeling depends on its use as a metabolic intermediate for other amino acids. For instance, it would be pointless to add specifically labeled ^{15}N glutamic acid since this is a precursor to most of the amide groups

in the amino acids. There are some cases, such as aspartic acid, for which dilution can be minimized by the use of strains of bacteria deficient in particular metabolic pathways.

More recently, the incorporation of deuterium into proteins has been used to improve the spectra of very large proteins or protein complexes (35 kD and larger) (Yamazaki et al., 1994). The advantage of deuterium labeling is that it greatly reduces the relaxation of ^{13}C nuclei due ^{1}H nuclei during multidimensional NMR experiments.

2D Heteronuclear Correlation Experiments

2D Heteronuclear correlation experiments correlate a heteronuclear resonance (^{15}N, ^{13}C etc.) with a proton resonance. In theory we could start and end on either nucleus as long as we passed through both. However, the sensitivity, S/N, of the correlation is governed by the following proportionality:

$$S/N \propto \gamma_a \gamma_b^{3/2} \left(1 - e^{-T_1/T_{rc}} \right) \tag{1.38}$$

where γ_a is the gyromagnetic ratio of the nucleus excited at the beginning of the experiment, γ_b is the gyromagnetic ratio of the detected nucleus, T_1 is the longitudinal relaxation time of the nucleus excited at the beginning of the experiment, and T_c is the recycle time of the experiment. Since the gyromagnetic ratio of ^{1}H is four times larger than that of ^{13}C and 10 times larger than that of ^{15}N, experiments are designed to begin and end on protons. The overall gain in sensitivity compared to experiments which start on proton and are detected on the heteronucleus is:

$$n\left(\gamma_{^1H} / \gamma_S \right)^{3/2} \tag{1.39}$$

where n is number of protons attached to the S nucleus, $\gamma_{^1H}$ is the gyromagnetic ratio of ^1H, and γ_S is the gyromagnetic ratio of the heteronucleus (^{13}C or ^{15}N). Thus, the gains in sensitivity are 31 fold for amide protons, 24 fold for methyl protons, 16 fold for methylene protons, and 8 fold for methine protons.

There are two basic types of heteronuclear correlation experiments. These are distinguished by the coherence order at which the transferred magnetization evolves during t_1 evolution. The HSQC (Heteronuclear Single Quantum Coherence) experiment uses single quantum coherence while the HMQC (Heteronuclear Multiple Quantum Coherence) experiment uses multiple quantum coherence. In the 2 dimensional form, these experiments provide similar information. That is, the spectrum consists of peaks which correspond to correlations between proton and the heteronuclei of interest. For instance, in the case of a ^{15}N HSQC or HMQC the spectrum consists of peaks for each amide nitrogen and proton pair in the protein. An example of such a spectrum is shown in Figure 2.8.

This type of spectrum is very useful during the assignment process. Although a 2D ^{15}N correlation experiment does not provide correlations that identify spin system types or make connections between spin systems, it does provide a master reference for all of the backbone amides in a protein. This is particularly important since almost all backbone assignment schemes use connectivities involving the amide proton.

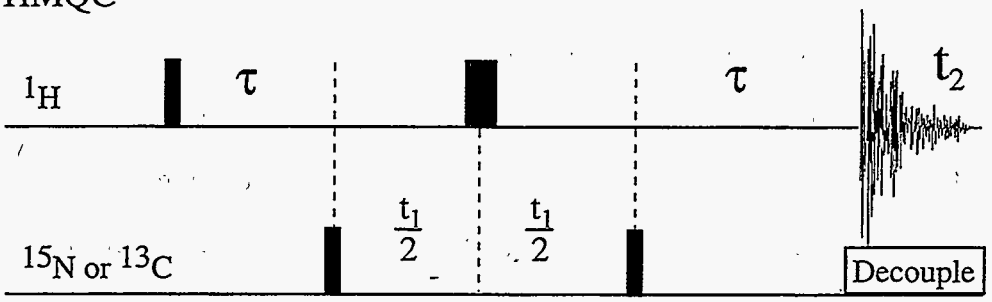
Furthermore, since these experiments provide probes (the amide nitrogen /amide proton correlation) at every residue except prolines and since these probes are very sensitive to conformational changes, these experiments can provide an excellent source of information for mapping conformational changes in a protein under various conditions. This type of strategy is pursued in chapter 5.

The HMQC Experiment

The HMQC pulse sequence is shown in Figure 1.7A. The product operator analysis of the HMQC is as follows:

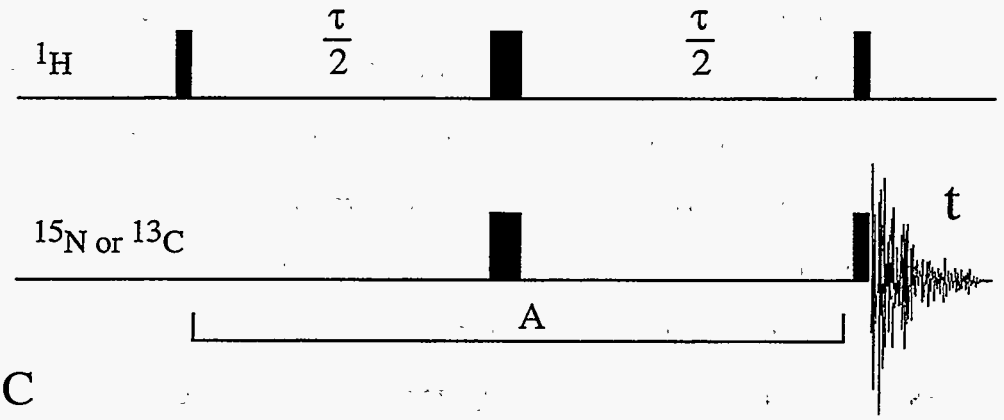
A

HMQC



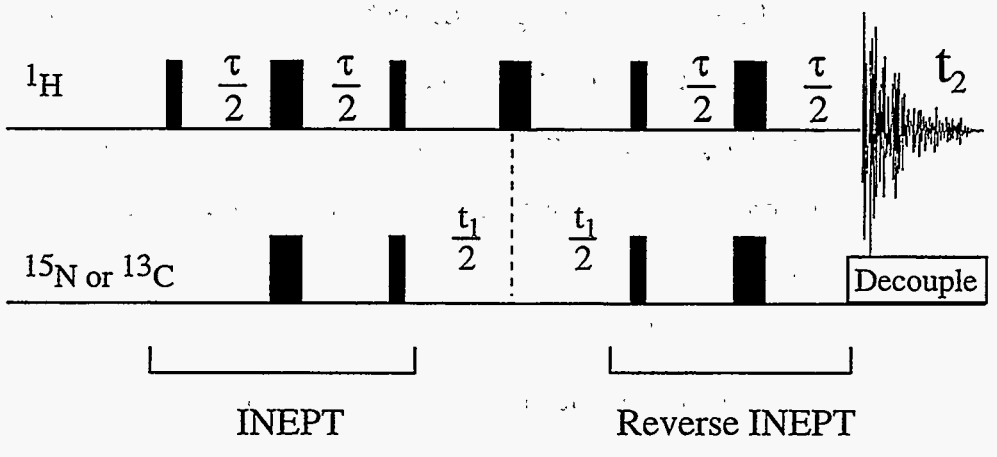
B

INEPT



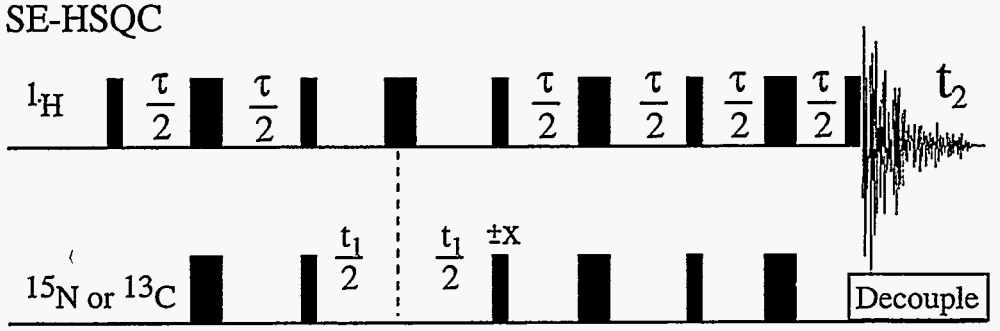
C

HSQC



D

SE-HSQC



E

Constant-time HSQC

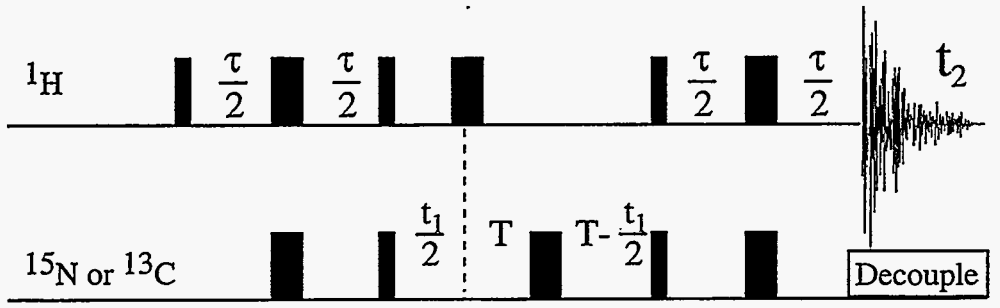


Figure 1.7: Pulse sequences of heteronuclear correlation experiments. The thin bars represent $\pi/2$ pulses and the thick bars represent π pulses. The details of these experiments are given in the text.

$$
\begin{aligned}
&\xrightarrow{\pi/2 I_x} \xrightarrow{J_{XH}\tau \hat{I}_z \hat{S}_z} \xrightarrow{\pi/2 S_x} \xrightarrow{(\omega_S t_1)\hat{S}_Z} \\
I_z &\longrightarrow I_y \longrightarrow -2I_xS_z \longrightarrow -2I_xS_y \longrightarrow -2I_xS_y\cos(\omega_s t_1) \\
&\xrightarrow{\pi/2 S_x} \xrightarrow{J_{XH}\tau \hat{I}_z \hat{S}_z} \xrightarrow{(\omega_I t_2)\hat{I}_Z} \\
&\longrightarrow -2I_xS_z\cos(\omega_s t_1) \longrightarrow -I_y\cos(\omega_s t_1) \longrightarrow \qquad (1.40) \\
&-I_y\cos(\omega_s t_1)\cos(\omega_I t_2) + I_x\cos(\omega_s t_1)\sin(\omega_I t_2)
\end{aligned}
$$

where I represents the ^1H nucleus, S represents the heteronucleus (^{15}N or ^{13}C), and J_{XH} is the scalar coupling constant between the ^1H and the heteronucleus. Note that to obtain the maximum signal, τ is set to $1/(2J_{XH})$ (this is sometimes set slightly shorter to account for relaxation). This experiment starts with a preparatory pulse on ^1H which creates transverse magnetization. This is converted to antiphase magnetization with respect to the heteronucleus during the period τ by the scalar coupling Hamiltonian. Multiple quantum magnetization (a mix of zero and double quantum) is created by a 90° pulse on the heteronucleus. The magnetization is labeled with the chemical shift of the heteronucleus during the time period t_1. The multiple quantum is converted back to proton magnetization, which is antiphase with respect to the heteronucleus, by the action of a second 90° pulse on the heteronucleus. The antiphase magnetization is converted back into transverse magnetization on the proton by the scalar coupling Hamiltonian during the second τ period. Finally, the magnetization is labeled with the proton's chemical shift during the detection period, t_2.

The chemical shift evolution of the ^1H nuclei is refocused for the entire (excluding t_2) experiment by the proton 180° pulse in the middle of the t_1 experiment. Note that the multiple quantum coherence does not evolve under the influence of the active scalar coupling Hamiltonian between the proton and heteronucleus. One drawback of this experiment is that passive couplings involving other protons are not refocused. Thus, each peak of the HMQC is really a multiplet which leads to some loss of resolution and signal.

The INEPT and HSQC Experiments

The INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) experiment is a 1D experiment designed to take advantage of the sensitivity gains from transfer of magnetization from proton to heteronucleus described above (Morris & Freeman, 1979). The pulse sequence for the INEPT is shown in Figure 1.7B. The product operator analysis of the INEPT is as follows:

$$I_z \xrightarrow{\pi/2 I_x} I_y \xrightarrow{A} -2I_xS_z \xrightarrow{\pi/2 S_x} -2I_xS_y \xrightarrow{\pi/2 I_x} -2I_zS_y \qquad (1.41)$$

where A is the set of Hamiltonians applied during the time period A indicated in Figure 1.7B. Overall, this pulse sequence generates heteronuclear magnetization antiphase with respect to proton which is converted into observable magnetization during the acquisition time. The time period A starts with proton magnetization in the transverse plane and ends with proton magnetization antiphase with respect to the heteronucleus. This is caused by evolution under the scalar coupling Hamiltonian between the proton and the heteronucleus. Note that chemical shift evolution for both the heteronucleus and the proton is refocused by the pair of 180° pulses in the middle of the time period A. The 90° pulses at the end of the sequence make the conversion between the two types of antiphase magnetization. With this pulse sequence in hand, we can now analyze the HSQC.

The HSQC pulse sequence is shown in Figure 1.7C. As indicated in the figure, this experiment consists of an INEPT experiment followed by a reverse INEPT experiment. The product operator analysis of the HSQC is as follows:

$$\begin{aligned} I_z &\xrightarrow{\text{INEPT}} -2I_zS_y \xrightarrow{(\omega_S t_1)\hat{S}_Z} -2I_zS_y\cos(\omega_s t_1) \xrightarrow{\pi/2 S_x\ \pi/2 I_x} \longrightarrow \qquad (1.42) \\ -2I_yS_z\cos(\omega_s t_1) &\xrightarrow{J_{XH}\tau\hat{I}_z\hat{S}_z} -2I_x\cos(\omega_s t_1) \xrightarrow{\text{Detect}} \end{aligned}$$

The first part of the HSQC is an INEPT which generates heteronuclear magnetization antiphase with respect to proton. This magnetization evolves under the chemical shift Hamiltonian of the heteronucleus during the time period, t_1, while evolution of the scalar coupling, J_{XH}, and the chemical shift of the proton are refocused by the 180° proton pulse. The magnetization is then converted back to proton magnetization antiphase with respect to the heteronucleus by 90° pulses on both nuclei. This then evolves back to transverse proton magnetization under the scalar coupling Hamiltonian. Once again, the chemical shift evolution of the heteronucleus and the proton is refocused by 180° pulses on both nuclei. The transverse magnetization is labeled with the chemical shift of the proton during the detection period, t_2.

This HSQC has significantly more radiofrequency pulses than the HMQC. Before the advent of modern spectrometers with precise control of pulse lengths and power, this fact made the HMQC a superior experiment. However, with the ability to precisely control pulses, the HSQC is generally now favored due to its superior relaxation and coupling properties (i.e. it does not experience passive couplings).

The Sensitivity Enhanced HSQC

The HSQC can be done in a sensitivity enhanced manner (Palmer et al., 1991). The pulse sequence for the sensitivity enhance HSQC is shown in Figure 1.7D. In the HSQC, the heteronuclear magnetization antiphase with respect to proton evolves under the heteronuclear chemical shift Hamiltonian to give two orthogonal terms $2I_zS_y$ and $2I_zS_x$ modulated by the cosine and sine of the chemical shift, respectively. In the unenhanced version of the HSQC only one of these terms, $2I_zS_x$, is refocused by the reverse INEPT and detected. The sensitivity enhanced HSQC refocuses both terms yielding a gain in signal to noise of up to $\sqrt{2}$.

As in the unenhanced HSQC, the experiment starts with an INEPT. Note that the phase of the second 90° pulse on the heteronucleus is phase cycled between $\pm x$. The product operator analysis of the sensitivity enhanced HSQC after the t_1 evolution period is

presented without the sine and cosine modulation terms for simplicity. Note that the upper line, A, is following one branch, I_zS_y, of the magnetization while the bottom line, B, follows the other branch, I_zS_x:

$$\begin{aligned} &\overset{\pi/2I_x\ \pi/2S_{\pm x}}{\longrightarrow\ \longrightarrow} \qquad \overset{J_{XH}\tau\hat{I}_z\hat{S}_z}{\longrightarrow} \\ A.\ I_zS_y &\longrightarrow \pm I_yS_z \longrightarrow \pm I_x \\ &\overset{\pi/2I_x\ \pi/2S_{\pm}}{\longrightarrow\ \longrightarrow} \qquad \overset{J_{XH}\tau\hat{I}_z\hat{S}_z}{\longrightarrow} \\ B.\ I_zS_x &\longrightarrow I_yS_x \longrightarrow I_yS_x \end{aligned} \tag{1.43}$$

Instead of detecting at this point, the I_x magnetization is placed along the z-axis for storage while the multiple quantum term, I_yS_x is converted into detectable magnetization.

$$\begin{aligned} &\overset{\pi/2I_y\ \pi/2S_y}{\longrightarrow\ \longrightarrow} \qquad \overset{J_{XH}\tau\hat{I}_z\hat{S}_z}{\longrightarrow} \qquad \overset{\pi/2I_x}{\longrightarrow} \\ A.\ \pm I_x &\longrightarrow \pm I_z \longrightarrow \pm I_z \longrightarrow \pm I_y \\ &\overset{\pi/2I_y\ \pi/2S_y}{\longrightarrow\ \longrightarrow} \qquad \overset{J_{XH}\tau\hat{I}_z\hat{S}_z}{\longrightarrow} \qquad \overset{\pi/2I_x}{\longrightarrow} \\ B.\ I_yS_x &\longrightarrow -I_yS_x \longrightarrow I_x \longrightarrow I_x \end{aligned} \tag{1.44}$$

Thus, we generate observable terms from both components present during t_1. The phase of the second heteronuclear 90° pulse is cycled between $\pm x$ to allow collection of absorption phased specta. When successive FID's are added and subtracted, they yield pure absorption spectra. The full sensitivity enhancement is obtained from adding these two spectra together.

In practice the full $\sqrt{2}$ gain in sensitivity is rarely achieved because of the extra radiofrequency pulses, which can suffer from inhomogeneity, and the extra time the magnetization spends relaxing in the xy plane. Nevertheless, this experiment can give impressive sensitivity gains for many systems.

There is a slightly improved version of the sensitivity enhanced HSQC called the PEP-Z HSQC (Akke et al., 1994). This experiment is based on the same principle and in

practice yield similar sensitivity gains. The PEP-Z HSQC was the basic 2D heteronuclear correlation experiment used for the analysis of the constitutive mutants of the receiver domain of NTRC in chapter 5.

The Constant Time HSQC

One important variant of the HSQC is the constant-time HSQC (Vuister & Bax, 1992). During the heteronuclear evolution period, t_1, scalar couplings between heteronuclei are not refocused. Thus, in the case of a ^{13}C HSQC, the very large 1 and 2 bond carbon-carbon couplings in the sidechains will cause undesirable multiplet peaks to appear. In order to achieve homonuclear broadband decoupling of ^{13}C during the t_1 period, a constant-time evolution period is used. This is shown in the pulse scheme for the constant-time HSQC in Figure 1.7E. Overall, the pulse scheme is similar to the conventional HSQC with an INEPT followed by an evolution period (constant-time) followed by a reverse INEPT and detection. During the constant time period, the J_{CH} coupling should be refocused, the ^{13}C chemical shift should evolve for a period t_1 and the J_{CC} coupling must be eliminated. This is accomplished by keeping the total evolution time period constant, but incrementing the time at which a ^{13}C 180° pulse appears. The time of evolution for each of the relevant Hamiltonians discussed above will be considered in turn. The J_{CH} coupling evolves for a period:

$$t_1/2 - T + \left(T - t_1/2\right) = 0 \qquad (1.45)$$

Note that the change in the sign of the evolution is due to refocusing of the scalar coupling by both the 180° proton pulse after the first $t_1/2$ period and by the ^{13}C 180° pulse after the first $t_1/2+T$ period. ^{13}C chemical shift evolves for a time period:

$$t_1/2 + T - \left(T - t_1/2\right) = t_1 \qquad (1.46)$$

Note that the ^{13}C chemical shift Hamiltonian is refocused by the incremented ^{13}C 180° pulse. The period of evolution for J_{CC} coupling is:

$$t_1/2 + T + \left(T - t_1/2\right) = 2T \qquad (1.47)$$

Thus, the J_{CC} coupling evolves for a constant time 2T and no modulation due to this coupling is observed. In order to maximize the signal from this experiment, the time period 2T is set to a multiple of $1/J_{CC}$. Scalar coupling between carbonyl carbons and aliphatic carbons is removed by the first selective 180° carbonyl pulse.

One disadvantage of the constant time HSQC is the lengthy time the magnetization must stay in the transverse plane during frequency labeling. This can lead to a loss of sensitivity due to relaxation. However, in the case of uniformly ^{13}C labeled protein samples, the gains from the homonuclear broadband decoupling usually outweigh the loss of sensitivity due to relaxation.

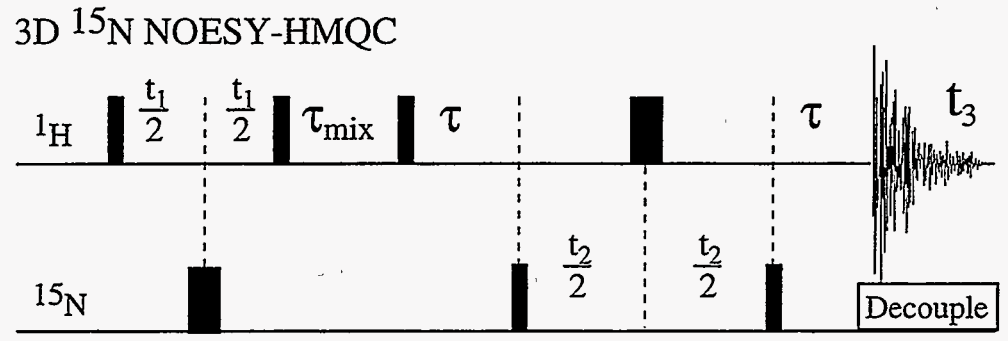
3 and 4 Dimensional Experiments

There are basically two classes of 3 and 4D experiments for backbone assignment (Bax & Grzesiek, 1993). The first class uses a heteronucleus (^{15}N or ^{13}C) to provide another chemical shift parameter with which to separate resonances. The second type, triple resonance experiments, use the heteronuclei to transfer magnetization through the backbone. In either case, however, these higher dimensional experiments are simply concatenations of the HMQC, INEPT/HSQC, NOESY, TOCSY and COSY experiments.

The ^{15}N-edited 3D NOESY and TOCSY Experiments

The ^{15}N-edited 3D NOESY-HMQC (Kay et al., 1989; Marion et al., 1989b) and the ^{15}N-edited 3D TOCSY-HMQC (Driscoll et al., 1990) are examples of the first type of 3D experiment (Figure 1.8). These experiments take a ^{1}H-^{1}H NOESY or TOCSY

A



B

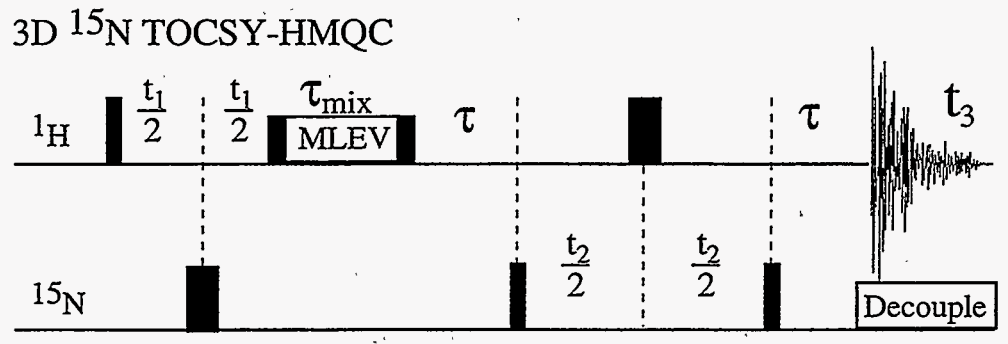


Figure 1.8: Pulse sequence of the 3D ^{15}N editedNOESY-HMQC (A) and the 3D ^{15}N edited TOCSY-HSQC(B). The tall, thin black bars represent $\pi/2$ pulses, the thick bars represent π pulses, and the short, thin bars represent trim pulses. The details of these experiments are given in the text.

spectrum, and spread out the resonances into a third dimension by adding a HMQC on to the end of the experiment.

Both of these experiments (Figure 1.8) work similarly. A preparation pulse on ^1H put magnetization into the xy plane. This magnetization is labeled with the ^1H chemical shift during the time period t_1. The J_{NH} coupling is removed by the 180° pulse in the middle of this period. The magnetization is then prepared for the mixing period by the next proton 90° pulse. The mixing period causes transfer of magnetization either through bond (TOCSY) or through space (NOESY). Note that the TOCSY mixing in Figure 1.8B locks the magnetization in the xy plane instead of along the z-axis as in Figure 1.2B. In this case, trim pulses are placed on either side of the mixing sequence to remove magnetization that is not aligned along the axis of the spin-lock. After the mixing period, an HMQC is performed. The magnetization is converted to multiple quantum magnetization of the ^1H and the ^{15}N. The ^{15}N chemical shift is labeled during the t_2 period and magnetization is converted back into observable transverse magnetization for the evolution of the amide proton chemical shift and detection during the time period t_3. Thus, these pulse sequences only detect magnetization that ends up on the amide proton.

The effect of these experiments is to take the amide region of a 2D ^1H-^1H NOESY or TOCSY and spread the spin systems into a third dimension based on the chemical shift of the amide nitrogen. This greatly reduces the overlap of the spectrum which facilitates backbone assignment.

The actual procedure of assignment with these experiments is exactly the same as their homonuclear counterparts. The same through space and through bond connectivities shown in Figure 1.6A are used. Indeed, one particularly convenient method of analyzing such 3D data is to create a 2D strip plot. This plot contains all of the spin systems arranged side by side. An example of a strip plot in sequential order can be seen in Figure 2.9. The initial assignment of the backbone of NTRC was carried with this type of analysis.

This type of analysis also provides a great deal of secondary structural information. The $d_{\alpha N}(i, i+3)$ and $d_{\alpha N}(i, i+4)$ connectivities are enough to form recognizable α helices in molecular dynamics simulations. The cross strand NOE's in regions of β sheet similarly yields recognizable secondary structure (see Figure 2.14). However, this type of information is insufficient to determine the structure of the secondary elements to high resolution. More importantly, this analysis fails to give information about how these elements are oriented with respect to each other in space. Thus, to determine a three dimensional structure, information from the sidechains is required.

Triple Resonance Experiments-the CBCA(CO)NH and the HNCACB

One of the weaknesses of the traditional assignment method described above is that it depends on through space coherences to make connectivities between amino acids in the protein sequence. These connectivities are not always unambiguous because through space interactions do not always arise from protons on sequential amino acids.

Triple resonance methods eliminate this ambiguity by relying exclusively on through bond coherences for backbone assignment (Ikura et al., 1990). These experiments work by exploiting the large 1 and 2 bond J couplings between heteronuclei on the peptide backbone of a uniformly ^{15}N, ^{13}C labeled protein. The relevant coupling constants are shown in Figure 1.6B.

There are an enormous number of different triple resonance experiments in the literature. However, only the CBCANH (Grzesiek & Bax, 1992b) and the CBCA(CO)NH (Grzesiek & Bax, 1992a) will be discussed. The pulse sequences and the magnetization pathways for these experiments are shown in Figure 1.9. This pair of experiments provides an extremely powerful method of sequentially assigning proteins.

The CBCA(CO)NH correlates the alpha and beta carbon resonances to the amide nitrogen and proton resonances of the next residue. The pulse sequence makes use of the large one bond couplings between ^{1}H and $^{13}C_{\alpha/\beta}$ (140 Hz), $^{13}C_{\alpha}$ and $^{13}C'$ (11 Hz), $^{13}C'$ and ^{15}N (15 Hz), and ^{15}N and ^{1}H (91 Hz). The experiment starts with an INEPT from the

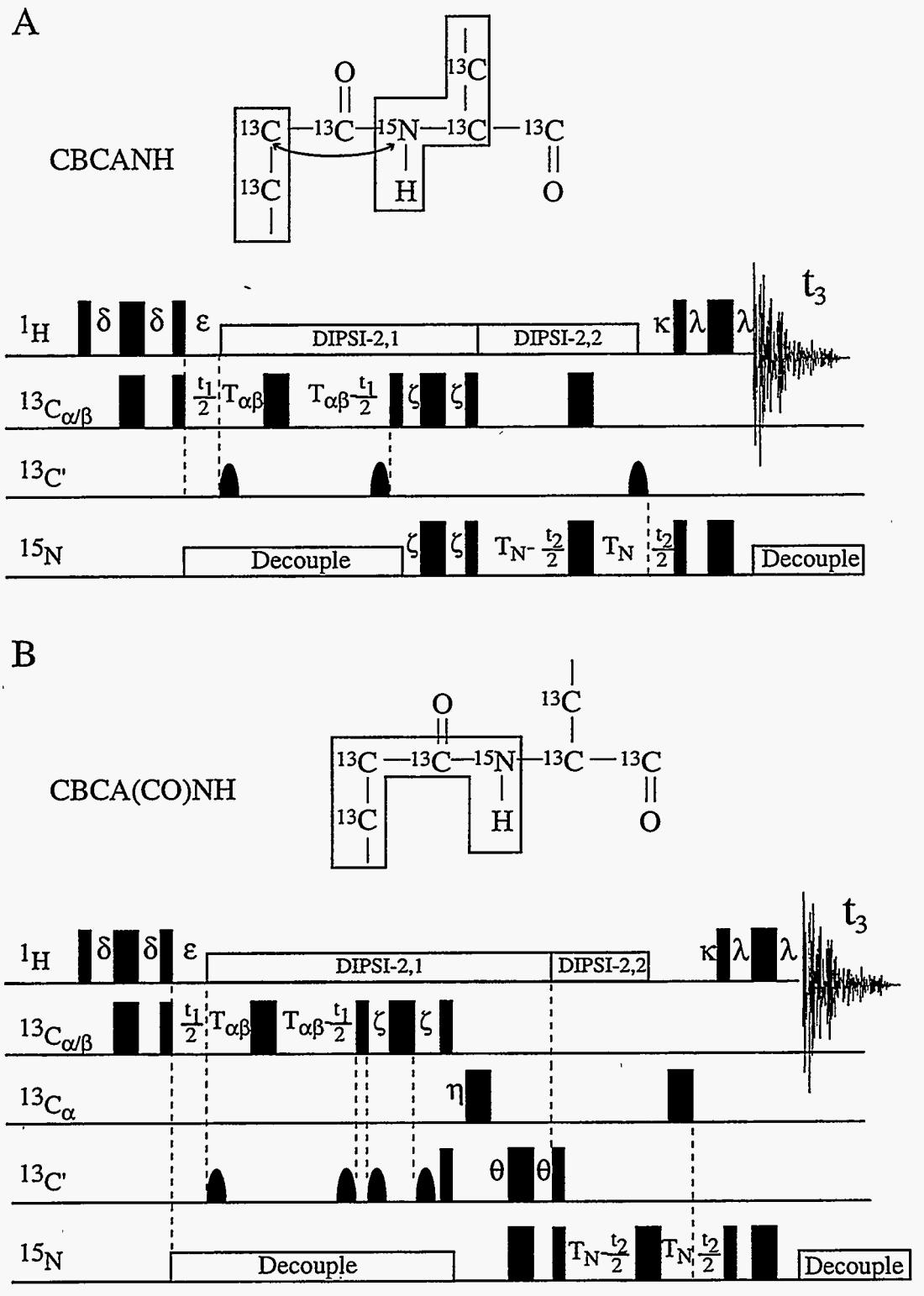


Figure 1.9: Schematic diagrams of the nuclei involved and the pulse sequences for the (A) CBCANH and the (B) CBCA(CO)NH. The delays δ, ε, κ, λ, and ζ are set to maximize various INEPT type transfers. Details are given in the text.

alpha and beta protons to their respective carbons. The carbons are frequency labeled during the first constant time period $2T_{\alpha/\beta}$. During this same period a fraction of the magnetization on C_β becomes antiphase with respect to C_α. This magnetization is converted to C_α magnetization antiphase with respect to C_β by the action of the 90° ^{13}C pulse at the end of the constant time period. At this point the magnetization which started on both H_α and H_β has been transferred to C_α and the carbon carrier frequency is set to the center of the C_α range rather than the center of the C_α/C_β range to take advantage of this fact. The magnetization is then passed to the carbonyl by another INEPT step. Note that carbonyl is treated as a separate spin system by using selective pulses that only excite the carbonyl carbons (the rounded pulses in Figure 1.9 on the carbonyl channel are indicative of this). Next, the magnetization is passed to the amide nitrogen by a third INEPT transfer. The amide nitrogen chemical shift is encoded during the t_2 time period. Finally, the magnetization is transferred to the amide proton by an INEPT for detection.

The CBCANH experiment correlates alpha and beta carbon resonances with the amide nitrogen and proton resonances of the same residue and the next residue. The pulse sequence works in a similar manner to that of the CBCA(CO)NH except that it relies on the direct couplings between the C_α and the amide nitrogen. Neither of these couplings is particularly large (11 Hz for the intraresidue and 7 Hz for the interresidue). Thus, although in theory this experiment contains all of the information needed for backbone assignment, in practice this is rarely the case because the interresidue connections are often missing.

However, the combination of the two experiments is particularly powerful. The CBCANH provides the intraresidue correlation for a particular amide while the CBCA(CO)NH provides the interresidue correlation for the same amide. Note that this can still leave the problem of determining the types of spin systems that are being linked together. This information can come from several sources. The 3D ^{15}N TOCSY-HMQC can classify amino acids as discussed above. The chemical shifts of the alpha and beta

carbons correspond well to amino acid type (Wishart et al., 1991). Finally, the 3D ^{13}C HCCH TOCSY (see next section) is a very powerful method of classifying amino acids.

In the case of NTRC, only the CBCA(CO)NH was acquired. This was sufficient to confirm the assignments made from the 3D 15N NOESY-HMQC and 3D 15N TOCSY-HMQC pair. A selection of strips from the CBCA(CO)NH of NTRC is shown in Figure 2.10.

Sidechain Assignment - the 3D ^{13}C HCCH-TOCSY

As already discussed, the chemical shift assignments of the sidechain carbons and protons are necessary for the determination of the 3 dimensional structure. The 3D ^{13}C HCCH TOCSY (Bax et al., 1990; Fesik et al., 1990) is a convenient experiment for this purpose. This experiment correlates all of the carbons and protons in the sidechain(with a few exceptions) via isotropic mixing of the aliphatic carbons. This takes advantage of the large (35 Hz) coupling between the carbons. The pulse sequence for the 3D ^{13}C HCCH TOCSY is shown in Figure 1.10A.

A preparation pulse places ^{1}H magnetization into the transverse plane. This magnetization is frequency labeled with the ^{1}H chemical shift during the t_1 time period. The magnetization is then passed to the aliphatic carbons by an INEPT. During the t_2 time period, the magnetization is labeled with the carbon chemical shift. A mixing period causes isotropic mixing which transfers magnetization to all of the other carbons in the spin system. The magnetization is then transferred back to the protons via a reverse INEPT for detection.

This experiment provides complete chemical shift assignments for the carbons and protons in most sidechains. Unfortunately, the aromatic sidechains are inaccessible by this method. The chemical shift difference between the aliphatic carbons and the aromatic carbons in an aromatic sidechain is too large to allow effective isotropically mixing. Links to these sidechains must be made from NOESY type information or a separate HCCH TOCSY optimized for the aromatic carbons.

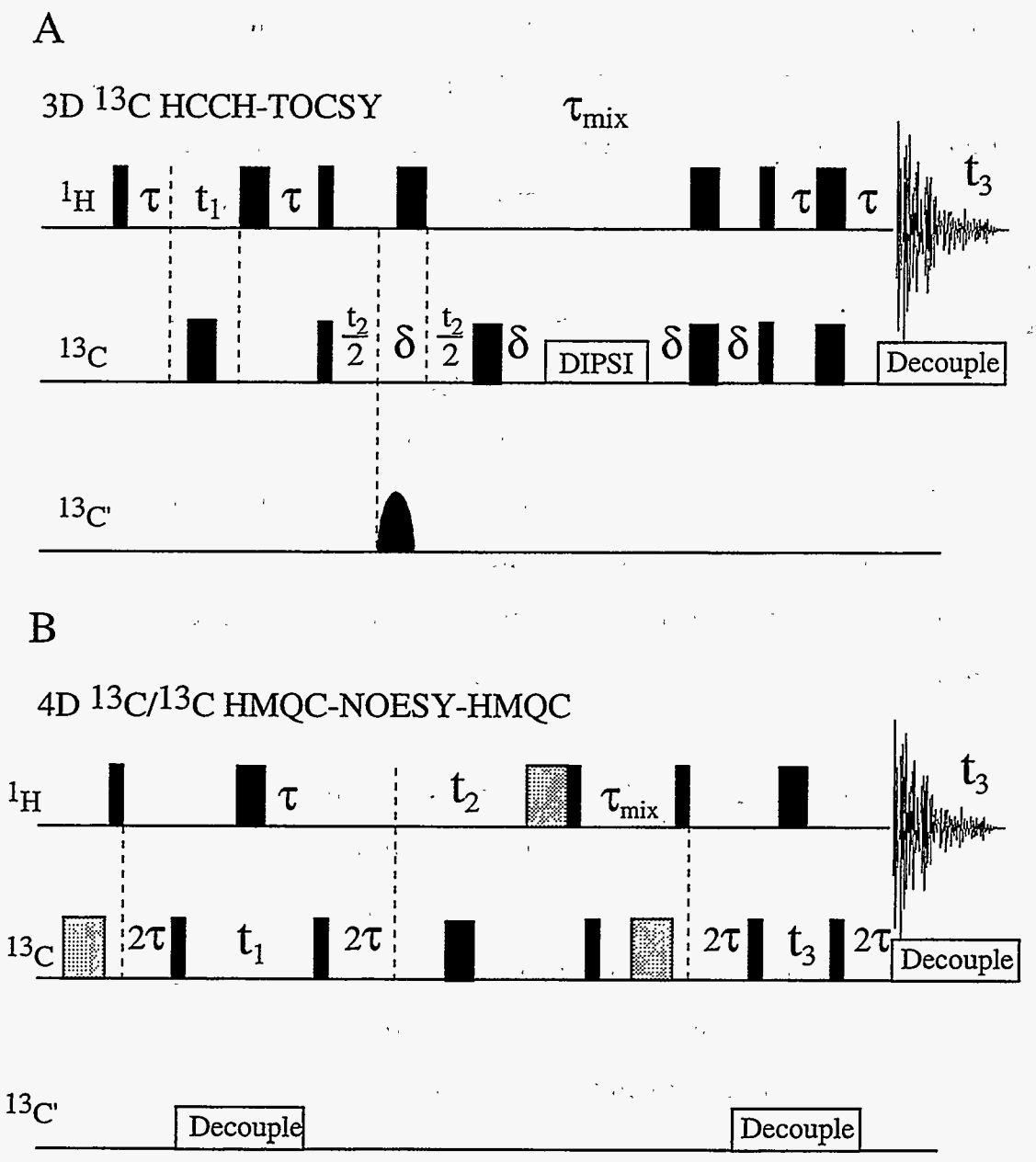


Figure 1.10: Pulse sequence of the 3D ^{13}C HCCH TOCSY(A) and the 4D $^{13}C/^{13}C$ NOESY-HMQC (B). The tall, thin black bars represent $\pi/2$ pulses and the thick bars represent π pulses. The thick grey bars in the 4D $^{13}C/^{13}C$ NOESY-HMQC experiment are purge pulses which remove unwanted magnetization.

This was the primary experiment used for the sidechain assignment of NTRC. A plane of this experiment, containing several complete spin systems, is shown in Figure 2.11. Note that this experiment is heavily aliased due to the extremely large carbon chemical shift range.

The 4D NOESY Experiment

The final experiment used for the determination of the intermediate resolution structure of NTRC was the 4D ^{13}C-^{13}C NOESY-HMQC (Clore et al., 1991). This experiment provides the long range distance information required to perform structure calculations. It correlates carbon proton pairs through space via a NOESY mixing period. Note that this is a 4D experiment because the chemical shift of each of the four nuclei in the two carbon proton pairs is labeled.

The pulse sequence for this experiment is shown in Figure 1.10B. An HMQC sequence first encodes the chemical shift of the first carbon proton chemical shifts. The magnetization is passed to another proton during a NOESY mixing period. Another HMQC is then performed to label second carbon proton pair.

This experiment is analyzed by picking the peaks in the four dimensional spectrum. Note that each peak has four chemical shift indices associated with it. The carbon chemical shifts are somewhat ambiguous due to the extreme aliasing required in the ^{13}C dimension to allow a reasonable acquisition time (96 hr).

Structure determination

The first step in structure determination (after sequence specific resonance assignment) consists of generating distance constraints from the 4D ^{13}C-^{13}C NOESY-HMQC data. This is done by comparing each peak of the NOESY (in the case of NTRC 1616 peaks) with the assignments generated from the 3D ^{13}C HCCH TOCSY. These assignments are made up by of pairs of protons and carbons. Thus, the process consists of trying to find possible matches among the TOCSY data for the starting and ending proton-carbon pairs of a NOESY peak. This is easily done by computer.

This analysis yields a great number of possibilities for each peak. Initially, rounds of structure calculations are performed with just unambiguous distance restraints. In the case of NTRC there were about 35 such long range restraints. This allowed a low resolution structure to be determined. This structure was used to resolve ambiguities in the remaining NOESY crosspeaks by ruling out possibilities that lie far outside the 5Å NOE distance limit. Thirty-five rounds of this type of refinement were performed to determine the intermediate resolution structure of NTRC. The final restraint file of contained 932 restraints. Note that this file contained restraints derived from a variety of sources. However, the bulk of the restraints come from the 4D ^{13}C-^{13}C NOESY-HMQC.

There are a number of prochiral groups in proteins. Without special methods to stereospecifically assign these groups, pseudoatom corrections must be added to the distance constraints involving them. This greatly reduces the precision of these restraints and degrades the overall quality of the structures. In particular, stereospecific assignment of the valine and leucine methyl pairs is important for the determination of high resolution structures (Guntert et al., 1989).

The structure calculations themselves were performed with the program X-PLOR (Brünger, 1992) using a hybrid distance geometry/simulated annealing protocol (Nilges et al., 1988). Distance geometry methods (Havel & Wüthrich, 1984) convert the list of restraints between atoms into a set of three dimensional coordinates for those atoms by using a triangle inequalities and an embedding procedure. These structures are then used as inputs for a simulated annealing molecular dynamics analysis. These method calculates the motions of atoms subject to a molecular force field and energy penalties for violations of the constraints (Brünger, 1992).

Families of structures consistent with the restraints are generated by this method. In the case of NTRC, typically families of 20-30 structures were determined. The comparison of these families yields a measure of the precision of the determined structures. The final family of structures for NTRC can be seen in Figure 2.15.

Chapter 2 The Three Dimensional Solution Structure of the Receiver Domain of NTRC

Introduction

Two Component Systems

In order to survive, all organisms must respond to changes in the environment. One of the most common mechanisms for this purpose in bacteria is the two component signal transduction systems (Parkinson & Kofoid, 1992). Members of this class have been found to play a role in sensing and responding to a wide variety of environmental stimuli such as nitrogen availability, osmolarity, and chemotactic information. Quite recently, such a system has been found as an essential component in the cell cycle of some bacteria (Quon et al., 1996). There are also examples of eukaryotic two component systems such as those involved in the ethylene receptor in *Arabidopsis thaliana* (Chang et al., 1993).

This family of signal transduction proteins was originally classified on the basis of sequence homology of two domains of about 250 and 130 amino acids (Nixon et al., 1986; Ronson et al., 1987). In general, the larger domain is a protein kinase which, using the γ-phosphate of ATP, autophosphorylates on a histidine residue in response to an environmental stimulus received by a receptor. The smaller component, the receiver domain or response regulator, then transfers the phosphate from the histidine to a carboxyl group on one of its own aspartic acids. This activates the receiver domain which transduces the signal to an attached domain or, in a few cases, a separate protein (Bourret et al., 1991). A general scheme for this process is shown in Figure 2.1.

Nitrogen Metabolism

Bacteria assimilate nitrogen into biomolecules primarily through incorporation of ammonia into the amino acids glutamate and glutamine (Woolfolk et al., 1966; Merrick & Edwards, 1995). While most of the nitrogen required for biosynthetic products by the cell, including the α-amino groups of the majority of the amino acids, is supplied by the amino group of glutamate, the γ-amido group of glutamine is also a major biosynthetic nitrogen

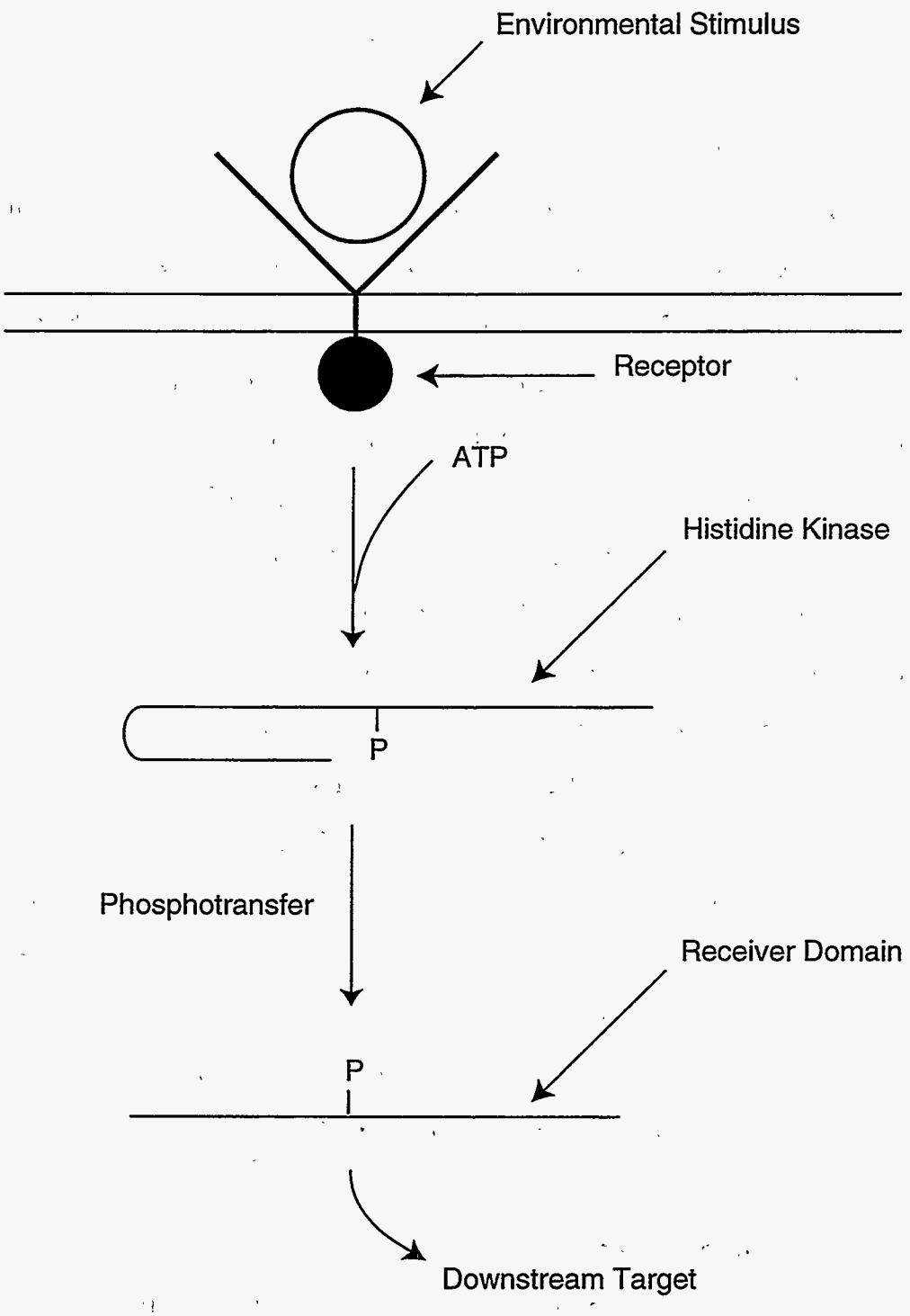


Figure 2.1: Schematic diagram of a typical two-component signal transduction system.

source. This group is used in the synthesis of tryptophan, asparagine, arginine, pyrimidines, purines and amino sugars. The major pathway to the formation of glutamine and glutamate requires two reactions. The first reaction, catalyzed by glutamine synthetase, incorporates ammonia into glutamine. This reaction is shown in Figure 2.2. The second reaction, catalyzed by glutamate synthase, is the reductive amination of α-ketoglutarate with glutamine as the nitrogen donor to form two glutamates. This reaction is shown in Figure 2.3. There is one other enzyme, L-glutamate dehydrogenase, that can incorporate ammonia into glutamate. This enzyme reductively aminates α-ketoglutarate to form glutamate. However, the K_m value for L-glutamate dehydrogenase is so high (~1mM) that this enzyme is not thought to make a huge contribution to ammonia assimilation (Sakamoto et al., 1975). Thus, the major pathway for the incorporation of nitrogen into the cell is a two step process involving glutamine synthetase and glutamate synthase.

As the first enzyme in this process, glutamine synthetase is an attractive target for regulation. It is regulated at both the transciptional and the post-translational levels. Post-translationally, it is controlled by allosteric regulation as well as covalent modification. There is a feedback inhibition of glutamine synthetase by tryptophan, histidine, CTP, AMP, carbamoyl phosphate, glucosamine-6-phosphate, glycine and alanine. The first six compounds are end products of glutamine biosynthesis while the alanine and glycine serve as monitors of overall cellular amino acid metabolism. Binding by any one of the compounds leads to only a slight inhibition of glutamine synthetase activity. However, binding by all eight products leads to a nearly complete shutdown of the enzyme (Hubbard & Stadtman, 1967; Woolfolk & Stadtman, 1967).

Another level of post-translational control - adenylylation of a particular tyrosine residue of glutamine synthetase - is overlaid on the cumulative feedback inhibition. This adenylylation renders glutamine synthetase more sensitive to feedback inhibition. Both the adenylylation and the deadenylylation are controlled by the enzyme adenylyl transferase. The activity of this enzyme is modulated by a regulatory protein, P_{II}, which binds to

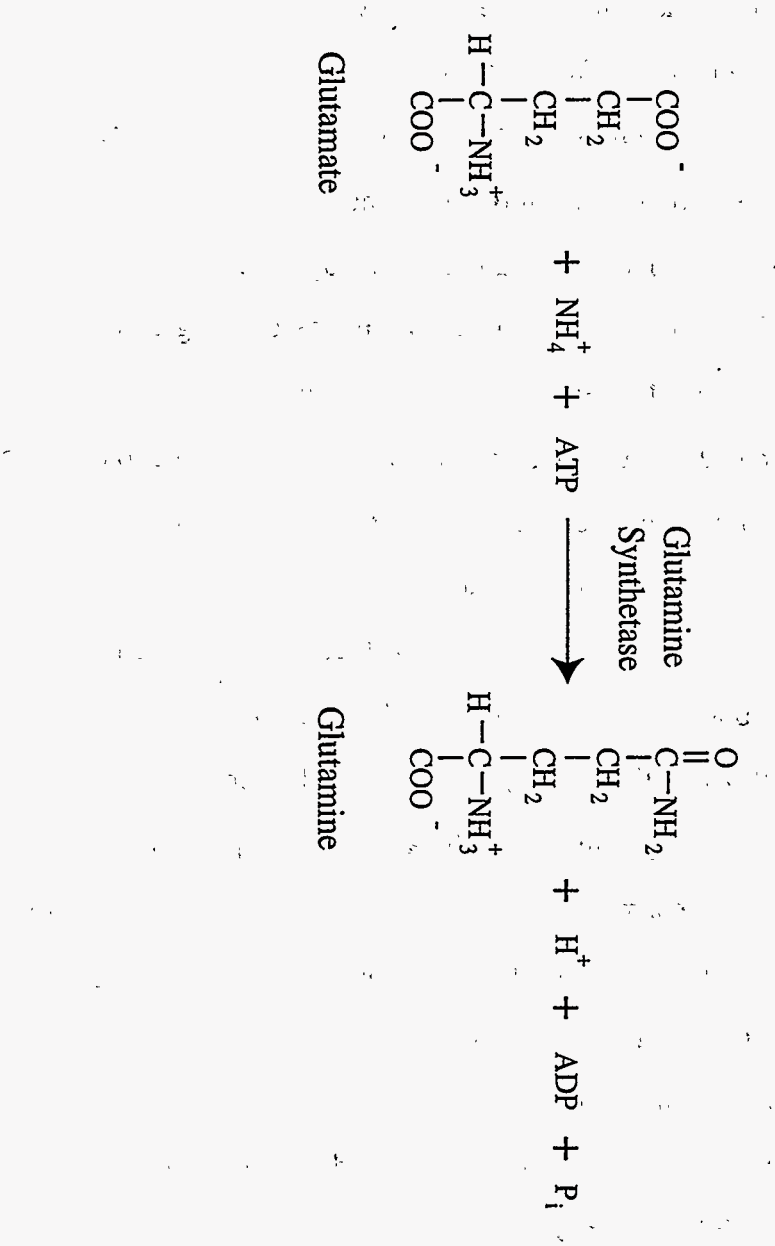


Figure 2.2: The reaction catalyzed by glutamine synthetase

Figure 2.3: The reaction catalyzed by glutamate synthase

adenylyl transferase (Son & Rhee, 1987). The effect of P_{II} on the activity of adenylyl transferase is determined, in turn, by a uridylylation of a tyrosine in P_{II}. Uridylylated P_{II} acts to stimulate the deadenylylation activity of adenylyl transferase which stimulates glutamine synthetase. Non-uridylylidated P_{II} acts to stimulate the adenylylation activity of adenyl transferase which inhibits glutamine synthetase. The uridylylidation state of P_{II} is controlled by the activity of the enzyme uridylyl transferase. The activity of uridylyl transferase inhibited by glutamine and stimulated by ATP and α-ketoglutarate (Chock et al., 1985; Rhee et al., 1989). A diagram of the overall control of the activity of glutamine synthetase is shown in Figure 2.4. When levels of glutamine are high, the activity of uridylyl transferase is inhibited which allows for the accumulation of non-uridylylated P_{II}. This causes adenylyl transferase to adenylylate glutamine synthetase which inhibits its activity. When levels of α-ketoglutarate are high, uridylyl transferase is stimulated to uridylylate P_{II}. This causes adenylyl transferase to deadenylylate glutamine synthetase which increases its activity. Therefore, the activity of glutamine synthetase is linked to the nitrogen state of the cell through the ratio of α-ketoglutarate to glutamine.

The levels of glutamine synthetase are also controlled at the transcriptional level in response to nitrogen availability (Stock et al., 1989b). Glutamine synthetase is transcribed from the *glnA* gene which resides in the *glnALG* operon. This promoter also contains the genes for two proteins, NTRC (NiTrogen Regulatory protein C) and NTRB (NiTrogen Regulatory protein B), which are involved in the regulation of glutamine synthetase. The *glnA* gene is controlled by two promoters *glnAp1* and *glnAp2*. Under conditions of nitrogen abundance, there is a low level of transcription from the *glnAp1* promoter, used by the σ^{70} form of RNA polymerase, which produces a small amount of glutamine synthetase. Under conditions of nitrogen scarcity, glutamine synthetase is produced in large amounts from the *glnAp2* promoter. Transcription from this promoter requires the products of the genes *ntrA* and *ntrC*. The protein product of *ntrA* , NTRA, is an alternate sigma subunit, σ^{54}, for RNA polymerase. σ^{54} guides RNA polymerase to different

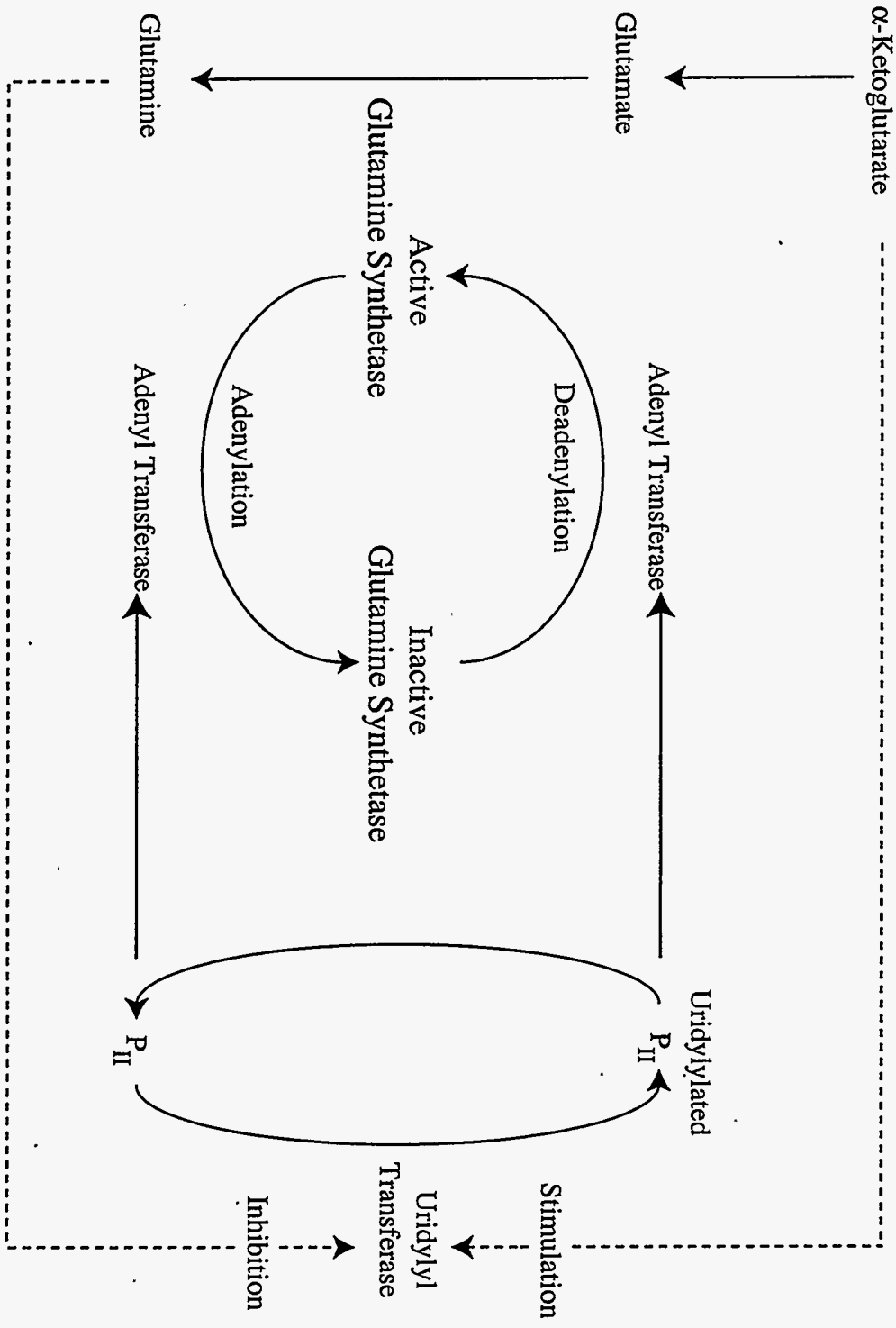


Figure 2.4: Diagram of the post-translational contol of the activity of glutamine synthetase.

promoter sequences than those preferred by the major sigma subunit, σ^{70} (Hirschman et al., 1985; Hunt & Magasanik, 1985). The protein product of *ntrC*, NTRC, is a transcriptional activator that binds to sites in the *glnAp2* promoter and stimulates the formation of an open complex between the DNA and the RNA polymerase at the start site (Popham et al., 1989; Wedel & Kustu, 1995).

The NTRC/NTRB two-component system

The NTRC protein is composed of three domains: an N-terminal two-component system receiver domain, a central activation domain with an ATPase, and a C-terminal DNA binding domain. The transcriptional activivty of NTRC is controlled via phosphorylation at aspartic acid 54 in the receiver domain (Sanders et al., 1992). The central domain ATPase is activated upon phsophorylation of the receiver domain and this ATPase activity is essential for transcription from the *glnAp2* promoter (Weiss et al., 1991). The activation of transcription depends on the formation of higher order oligomers of NTRC. The determinants for this oligomerization lie in the central domain of NTRC (Flashner et al., 1995). Unlike phosphorylation, deletion of the receiver domain fails to activate the central domain ATPase indicating that the N-terminal domain of NTRC actively stimulates the central domain in the phosphorylated state rather than being a repressor whose action is relieved by phosphorylation (Drummond et al., 1990; Weiss et al., 1992).

NTRC is phosphorylated via a phosphotransfer event from a histidine in the protein NTRB which is the histidine kinase of the NTRB/NTRC two componenet system (Ninfa & Magasanik, 1986; Keener & Kustu, 1988; Weiss & Magasanik, 1988). The activity of NTRC is controlled by NTRB. NTRB is not only a phosphodonor for NTRC, but also has a phosphatase activity which dephosphorylates NTRC. This phosphatase activity is stimulated by non-uridylylated P_{II} (Bourret et al., 1991). As discussed earlier, the uridylylation state of P_{II} is determined by the balance between glutamine and α-ketoglutarate. Therefore, the phosphorylation state of NTRC is linked to the nitrogen balance of the cell as indicated by the amounts of glutamine and α-ketoglutarate. When the

levels of free ammonia are low, NTRC accumulates in the phosphorylated form and the levels of glutamine synthetase are increased to scavenge the scarce amounts of ammonia. This process is shown in Figure 2.5.

NTRC and CheY

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., 1989a; Volz & Matsumura, 1991; Stock et al., 1993). 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 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.

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

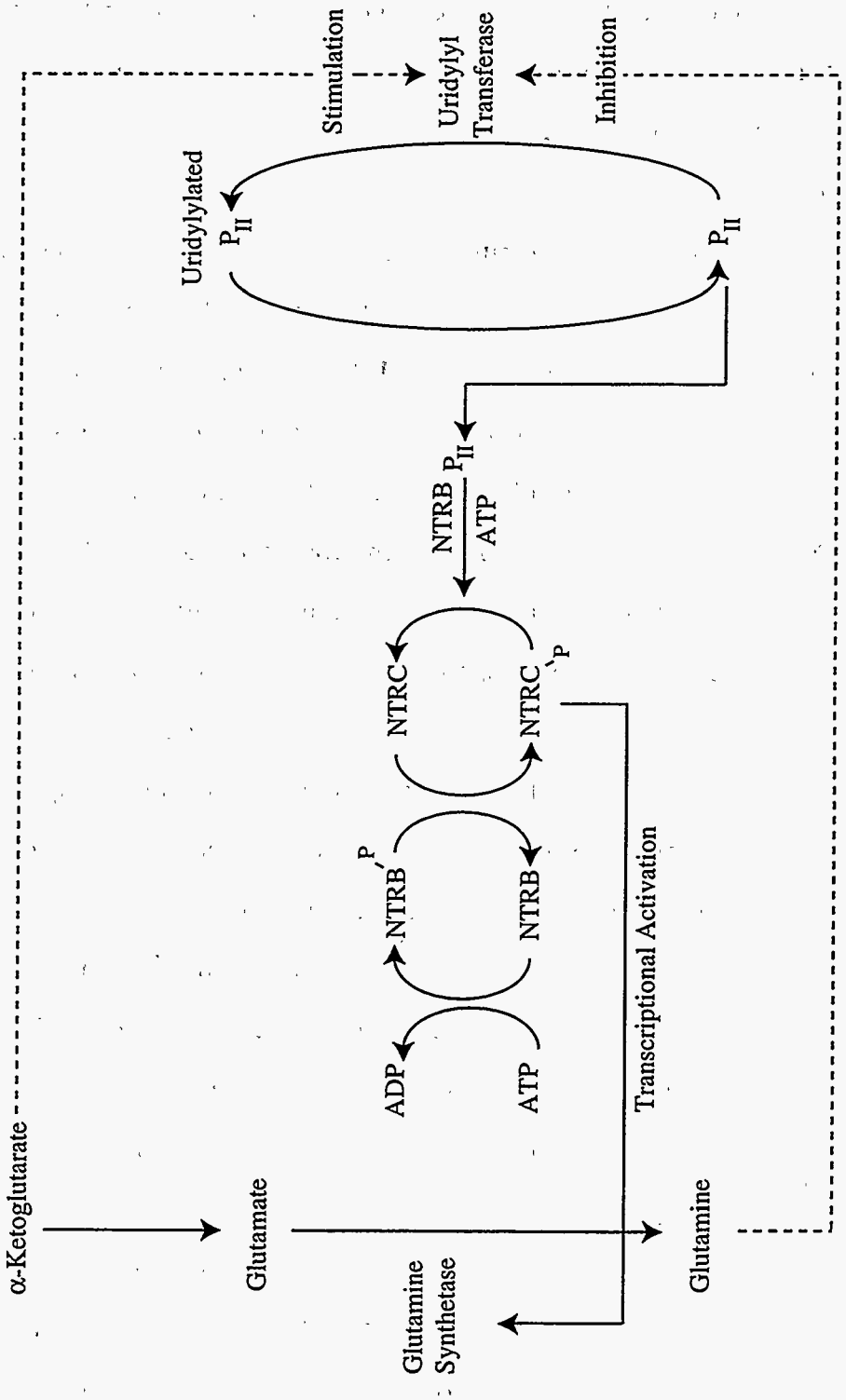


Figure 2.5: Regulation of the transcriptional activity of NTRC via phosphorylation. The phosphorylation state of NTRC is ultimately determined by the nitrogen balance of the cell as determined by the ratio of glutamine to α-ketoglutarate.

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 sodium 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, gel electrophoresis, 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 mM NTRC receiver domain 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 10 mM 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 0.1 M 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 and UV absorption at 280nm (NTRC extinction coefficient - 14060 M^{-1} cm^{-1}) (Gill & von Hippel, 1989).

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.

1D ^1H spectra were taken with a spectral width of 6944 Hz, 8192 total points, and the ^1H carrier placed on the H_2O resonance at 4.78 ppm. 2D TOCSY (Bax & Davis, 1985) and 2D NOESY(Macura et al., 1981) experiments were collected with spectral widths of 6944 Hz in both dimensions. The ^1H carrier was placed on the H_2O resonance at 4.78 ppm. The TOCSY mixing time was 64.6 ms, and the NOESY mixing time was 100 ms. A total of 1024 x 512 points were collected in the t_1 and t_2 dimensions. Data were apodized in each dimension with a sine-bell shifted 75° and skewed 1 in t_2 and 0.7 in t_2. Data were zero-filled to yield a 1024 x 1024 matrix upon Fourier transformation.

^{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 4.78 ppm, and the ^{15}N carrier set to 119.1 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 t_1, t_2, and t_3 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 t_1 (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 t_1 noise reduction routine, RT_1, to attenuate the streaking of intense methyl and aromatic signals, which tended to obscure weaker peaks (Manoleras & Norton, 1992). Post-processing with ANI and RT_1 used a threshold value, T, of 5 and a spread value, h, 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 t_1 ^{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 (t_1 and t_3), 56 points in the t_2 dimension, and 256 points in the t_4 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 a Fourier transformation of the t_3 ^{13}C dimension without apodization. The t_1 ^{13}C dimension was then extended to 12 complex points with linear prediction, apodized and Fourier transformed, followed by inverse transformation of the t_3 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 t_1 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-^1H 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

Initial Characterization

The initial characterization of the N-terminal domain of NTRC (all references to NTRC after this point refer to the N-terminal domain unless otherwise stated) was carried out with an unlabeled sample provided by Nancy Amy in Sydney Kustu's laboratory. Figure 2.6 shows a 1D ^1H spectrum of the amide region of NTRC. This spectrum

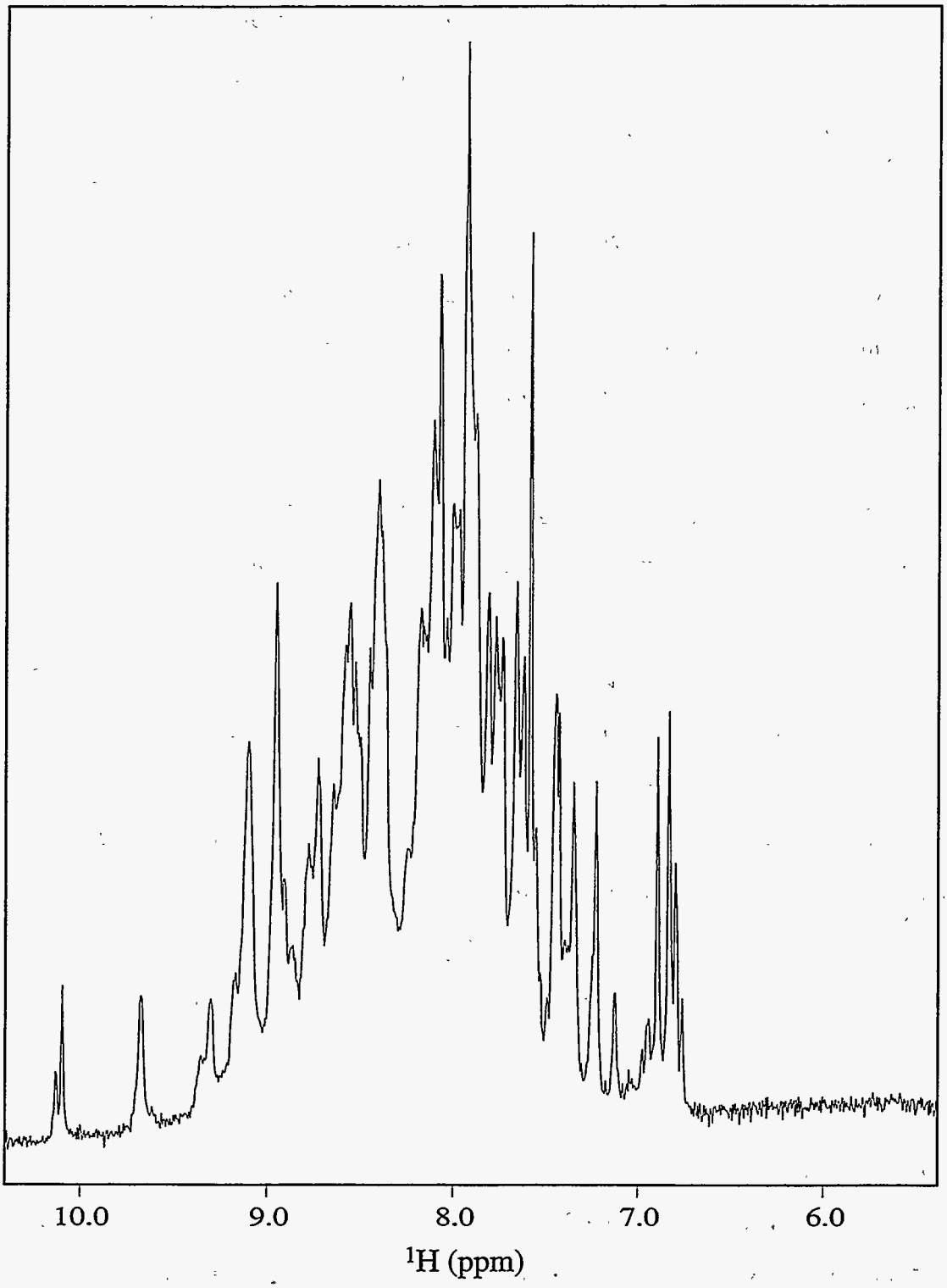


Figure 2.6: 1D ^{1}H spectrum of the amide region of the N-terminal receiver domain of NTRC at 600 MHz, pH 6.4, and 25° C.

generally shows the dispersion expected of a folded protein of this size. The linewidths were sharp enough to merit further analysis. A series of 1D spectra were taken to determine the optimal pH and temperature conditions. 25°C was chosen as the optimal temperature based on linewidth. The optimal pH was determined to be 6.4. The N-terminal NTRC domain precipitates at pH values below 6.0 (pI ~ 5.9). This is unfortunate because lower pH values minimize the rate of amide proton exchange with the solvent. The rapid exchange of amide protons can result in the loss of amide signals in the spectrum (Erikson et al., 1995).

Resonance Assignments

Initially, a set of homonuclear 2D experiments were taken. The 2D NOESY is shown in Figure 2.7. This experiment shows a large number of through space correlations indicative of a well folded protein. The 2D TOCSY experiment proved to be quite useful in the assignment process because, unlike the ^{15}N 3D TOCSY-HMQC, it contained nearly complete sidechain information for most spin systems. It was possible to extract almost complete information for well-seperated spin systems.

Figure 2.8 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 ^{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 $H\alpha$-NH, NH-NH, and $H\beta$-NH NOEs for assignment of residues in a traditional manner (Wüthrich, 1986). Figure 2.9 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 $H\alpha$ chemical shift dispersion which simplified much of the sequential assignment is also visible in Figure 2.9, with crosspeaks from W7 $H\alpha$ (6.55

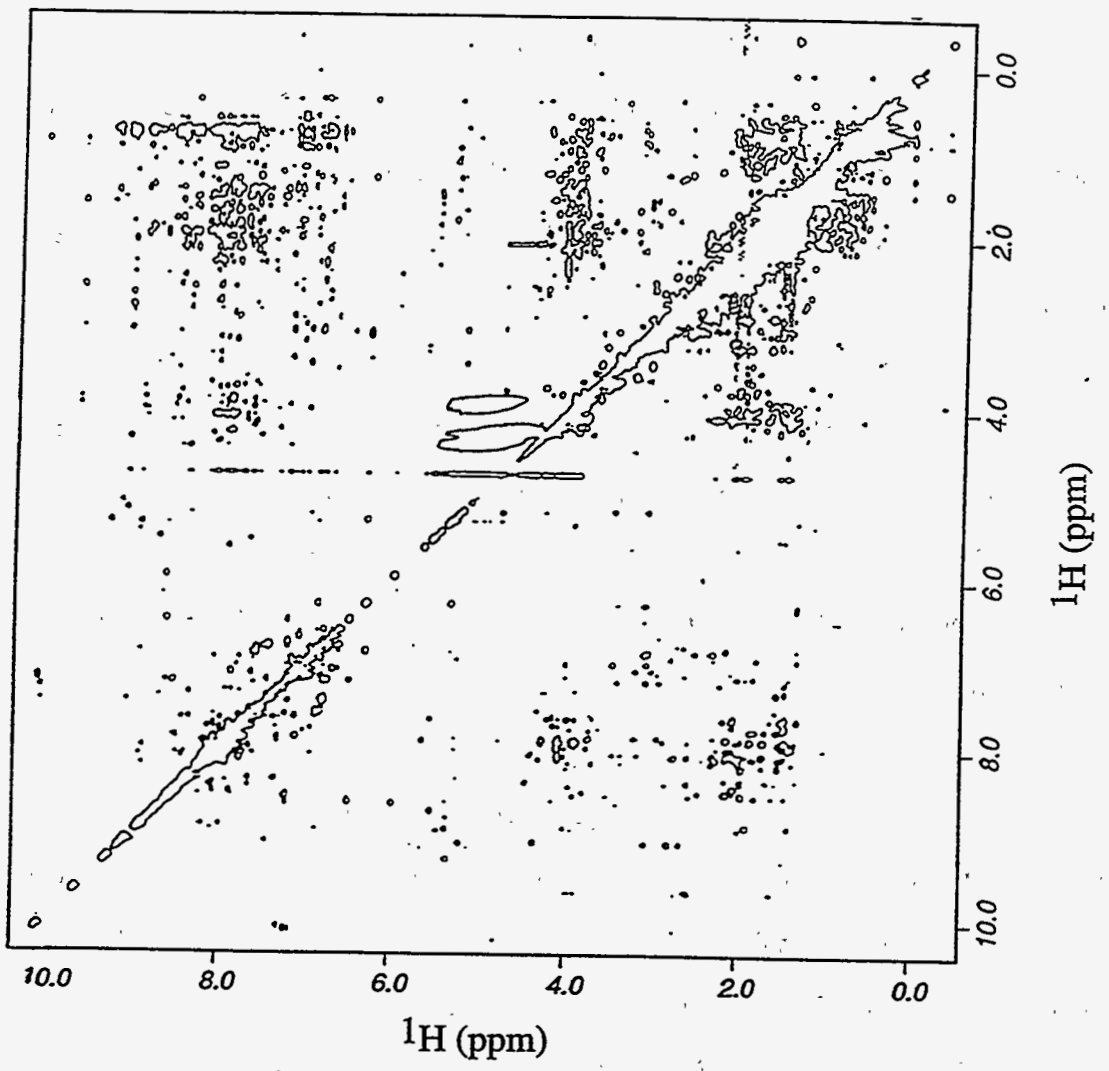


Figure 2.7: 2D ^{1}H NOESY spectrum of the N-terminal receiver domain of NTRC in 10 mM phosphate buffer, pH 6.4, 25 C°, and a 100 ms mixing time. Crosspeaks represent correlations between protons that are less than 5Å apart. The severe overlap of the crosspeaks in the upfield region makes unambiguous assignment of proton correlations difficult.

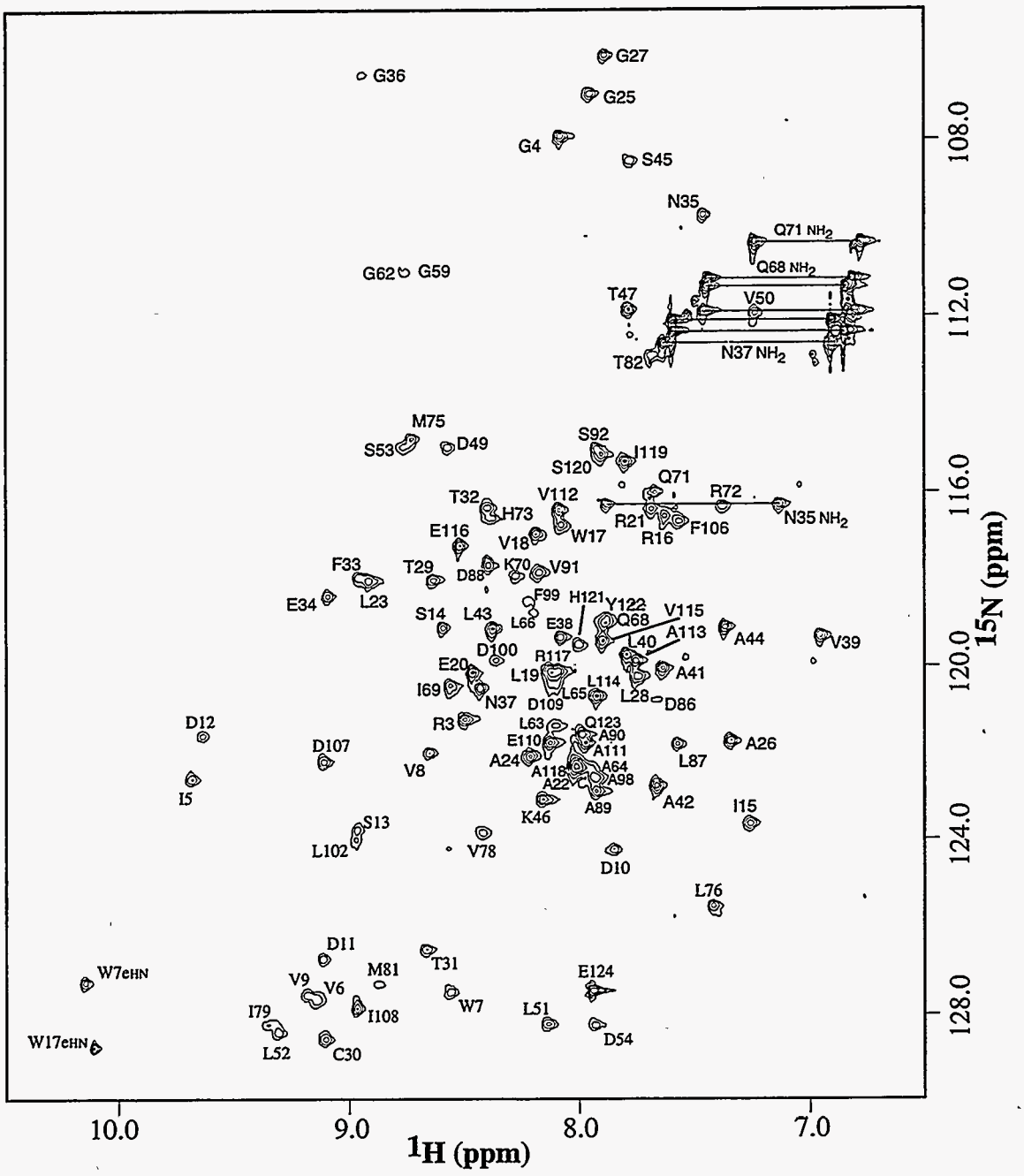


Figure 2.8: ^{1}H-^{15}N HSQC spectrum of uniformly ^{15}N-labeled NTRC receiver domain. Cross-peak assignments are indicated with the one-letter amino acid code and residue number.

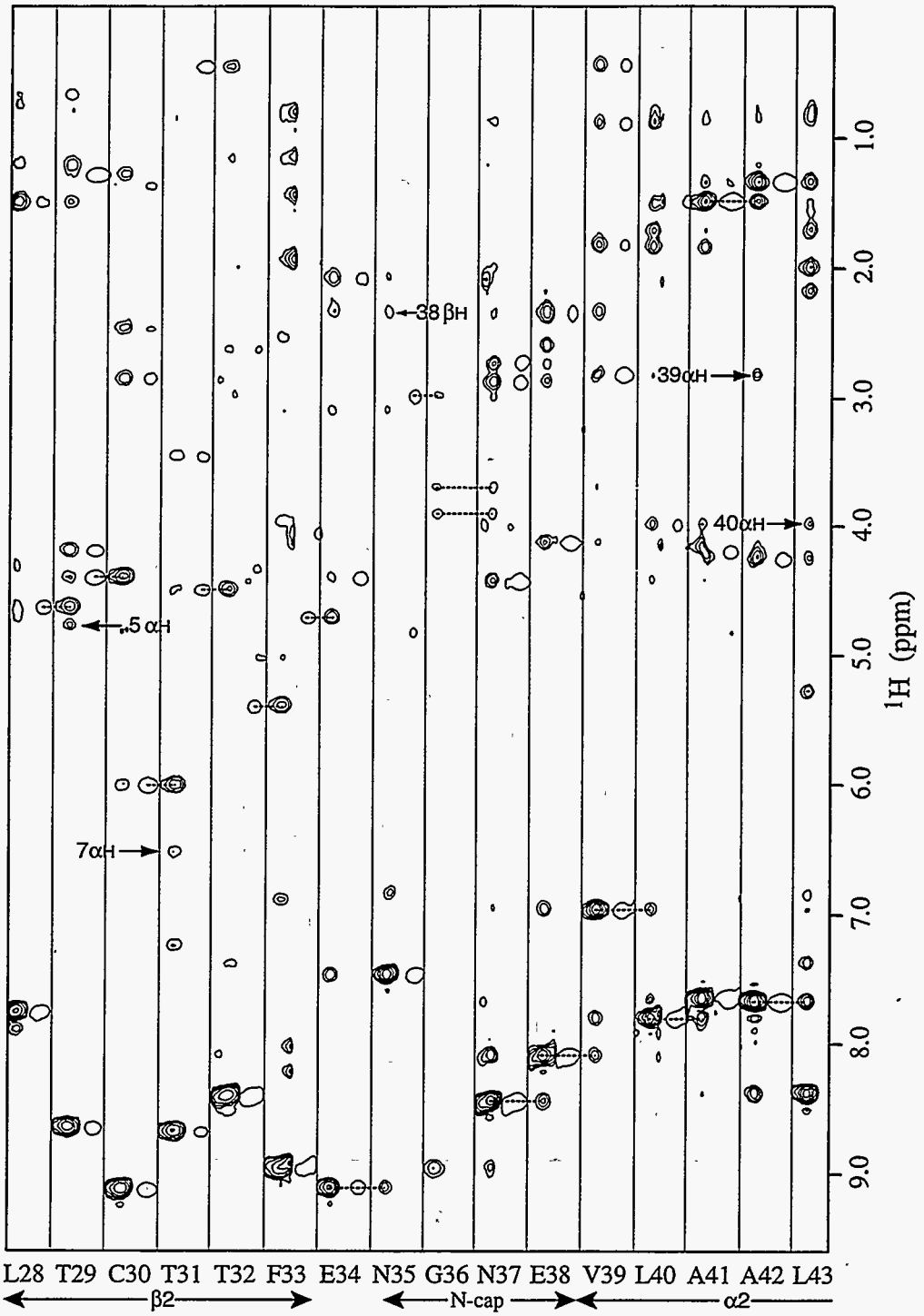


Figure 2.9: Strip plot of selected regions of the 3D ^{15}N NOESY-HMQC and TOCSY HMQC spectra of NTRC illustrating sequential resonance assignments for residues 28-43. NOESY and TOCSY strips are alternated with NOESY peaks displayed in multiple countours and the TOCSY peaks displayed with a single contour level. Sequential connectivities are indicated with dashed lines between NOESYand TOCSY peaks of sequential residues. Crosspeaks form 5 to 29 and 7 to 31 are due to cross-strand contacts in the β-sheet structure.

ppm) and V39 H_α (2.84 ppm) indicated. A number of assignments were facilitated 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 H_α and sidechain protons were assigned from the ^{15}N 3D TOCSY-HMQC spectrum.

The CBCA(CO)NH experiment was used to confirm the backbone assignments made from the ^{15}N 3D NOESY-HMQC and ^{15}N 3D TOCSY-HMQC spectra. Figure 2.10 shows a strip plot of the CBCA(CO)NH for residues 22 to 34. This experiment correlates the amide nitrogen and proton of residue i with C_α and C_β of the previous residue, i-1. Thus, this experiment alone is not sufficient to assign the backbone of NTRC. Usually the CBCA(CO)NH is interpreted in combination with an experiment, such as the CBCANH (Grzesiek & Bax, 1992b), that correlates the amide of residue i with C_α and C_β of the same residue. In this case the CBCA(CO)NH was used as a final confirmation of the backbone assignments in combination with the ^{15}N 3D TOCSY-HMQC and ^{13}C 3D HCCH-TOCSY.

One plane of the ^{13}C 3D HCCH-TOCSY experiment is shown in Figure 2.11. 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 H_α 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 determined

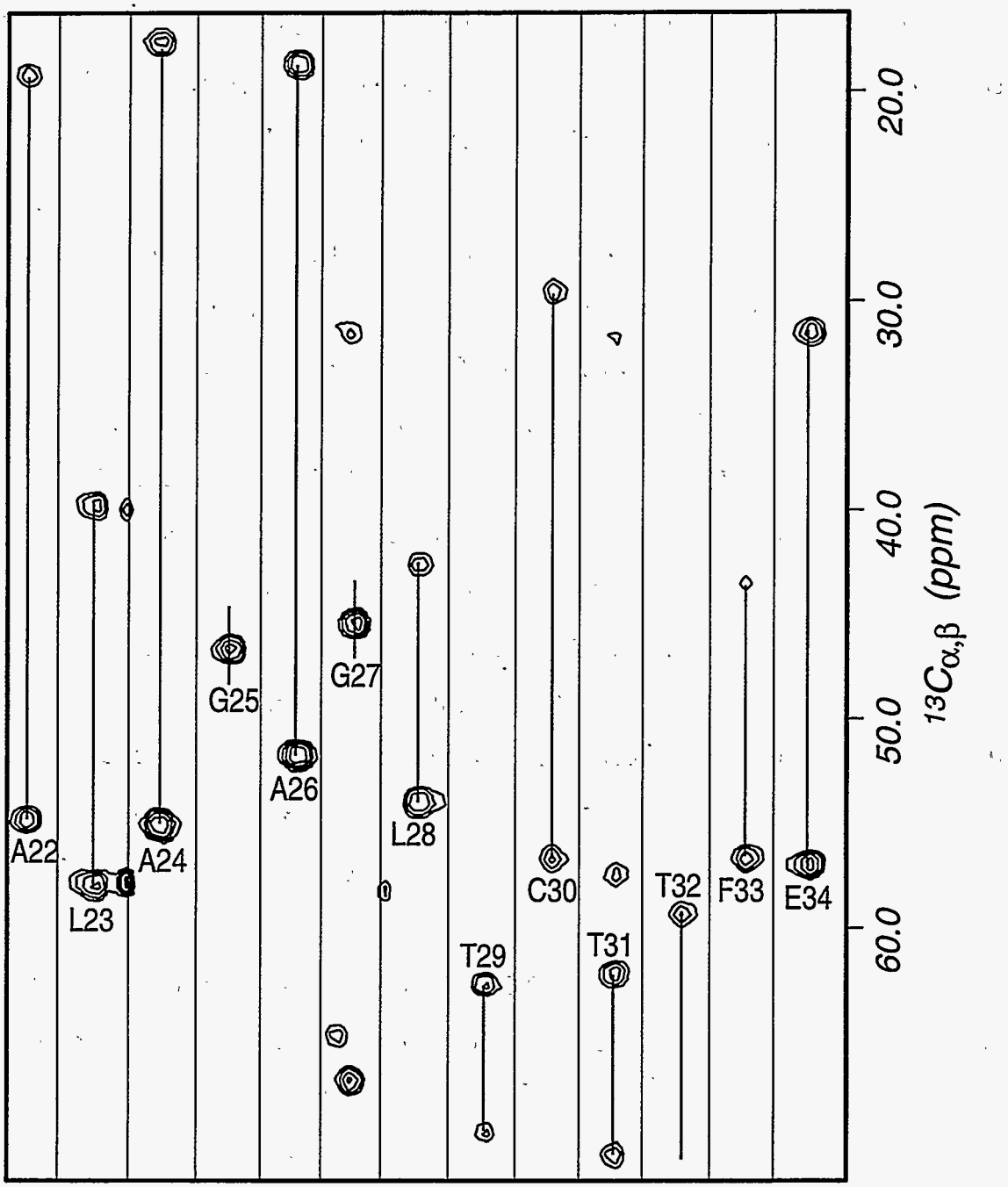


Figure 2.10: Strip plot of selected regions from CBCA(CO)NH spectrum of the NTRC receiver domain. Peaks correlate $^{13}C(\alpha, \beta)$ chemical shifts with the ^{15}N and ^{1}H chemical shifts of the following residue. The $^{13}C_{\alpha}$ correlations are labeled with residue type and number, and 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.

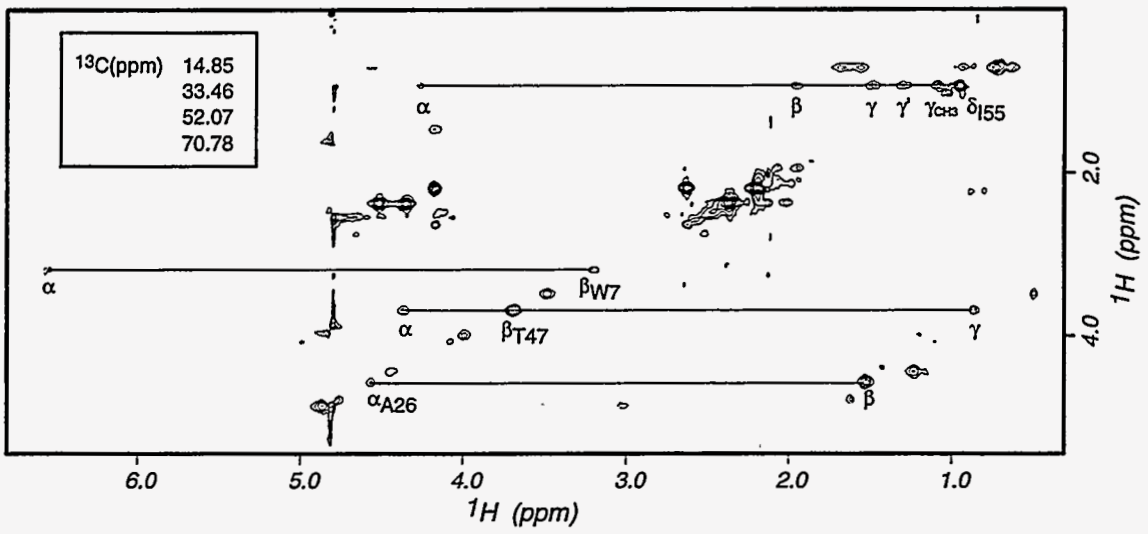


Figure 2.11: 2D ^{1}H-^{1}H plane of a 3D ^{13}C HCCH-TOCSY. Spin systems are labeled with residue assignment at the diagonal peak, with crosspeaks labeled by proton type. Due to extensive aliasing in the ^{13}C dimension, the $^{13}C\delta$ of I55 at 14.85 ppm, the $^{13}C\beta$ of W7 at 33.46 ppm and the $^{13}C\alpha$ of A26 at 52.07 ppm appear in the same plane as $^{13}C\alpha$ 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 2.1: ^{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 2.1 (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
C30	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
Q2			55.57	4.51	30.06	2.12	2.01	33.82	2.36	2.36			
Q68	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
Q71	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
Q95			58.21	4.12	28.13	2.17	2.17	34.02	2.45	2.45			
Q96			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
E20	120.15	8.46	60.74	4.03	29.19	2.12	2.12	36.41	2.27	2.22
E34	118.38	9.10	56.79	4.44	31.55	2.12	2.12		2.26	2.37
E38	119.34	8.08	58.82	4.15	30.71	2.36	2.36	36.47	2.63	2.41
E110	121.78	8.12	58.77	4.13	29.41	2.24	2.19	36.56	2.44	2.31
E116	117.24	8.52	59.89	3.87	28.70	2.16	2.04	35.92	2.55	2.16
E124	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
G4	107.92	8.08	45.38	4.02	3.95
G25	106.95	7.95	46.52	4.02	3.91
G27	106.06	7.88	45.29	4.36	3.80
G36	106.51	8.94	47.05	3.93	3.73
G59	110.99	8.76	45.53	4.12	3.87
G62	110.99	8.75	47.15	4.17	3.88
G97			45.38	4.03	4.03

His	N	HN	CA	HA	CB	HB1	HB2	CD2	HD2	CE1	HE1
H73	116.61	8.37	53.19	5.04	30.18	3.02	3.02	121.60	6.84	139.79	7.95
H84			59.39	4.38	29.37	3.40	3.34	120.54	7.31	138.30	8.31
H121	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*
I5	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
I15	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
I55			61.70	4.26	39.29	1.95	28.29	1.49	1.28	18.54	1.07	14.53	0.94
I69	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
I79	128.25	9.34	59.60	4.58	40.26	1.69				18.46	0.62	14.65	0.72
I80													
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 2.1 (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 2.1. (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

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 $H\alpha$-$H\alpha$ or $H\alpha$-$H\delta$ 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 NOE's 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 2.1 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 2.12 presents a summary of sequential and medium-range NOE information, as well as $J^3_{H\alpha\text{-}NH}$ coupling constant data and amide exchange information. Patterns of sequential Hβ-NH and NH-NH NOEs as well as Hα-NH(i, i+3) and Hα-Hβ(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 Hα-NH(i, i+3) NOEs for the start of helix 2 are highlighted in Figure 2.9. Strong sequential Hα-NH NOE's 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 2.13 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 2.12, 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

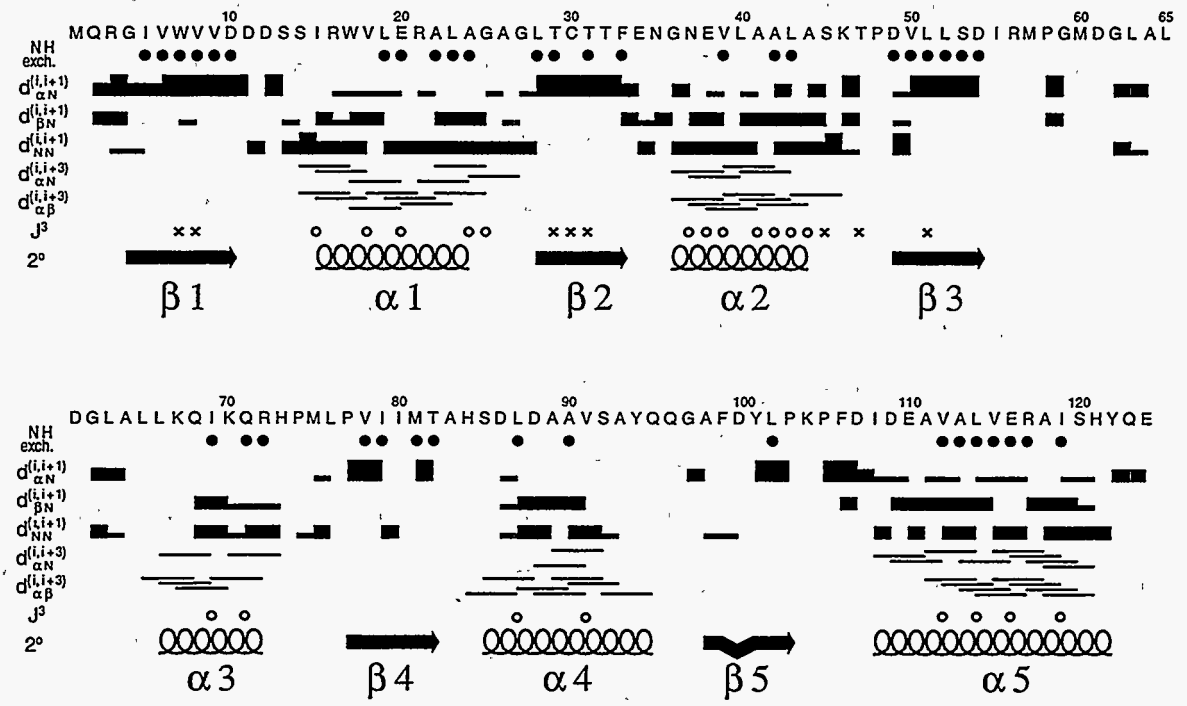


Figure 2.12: Summary of sequential and medium-range NOEs, slowly exchanging amide protons, and $J^3_{\alpha H\text{-}NH}$ values for NTRC receiver domain. Relative intensity of sequential NOEs is indicated by height of connecting box. Horizontal lines represent helical medium-range NOEs. Filled circles 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.

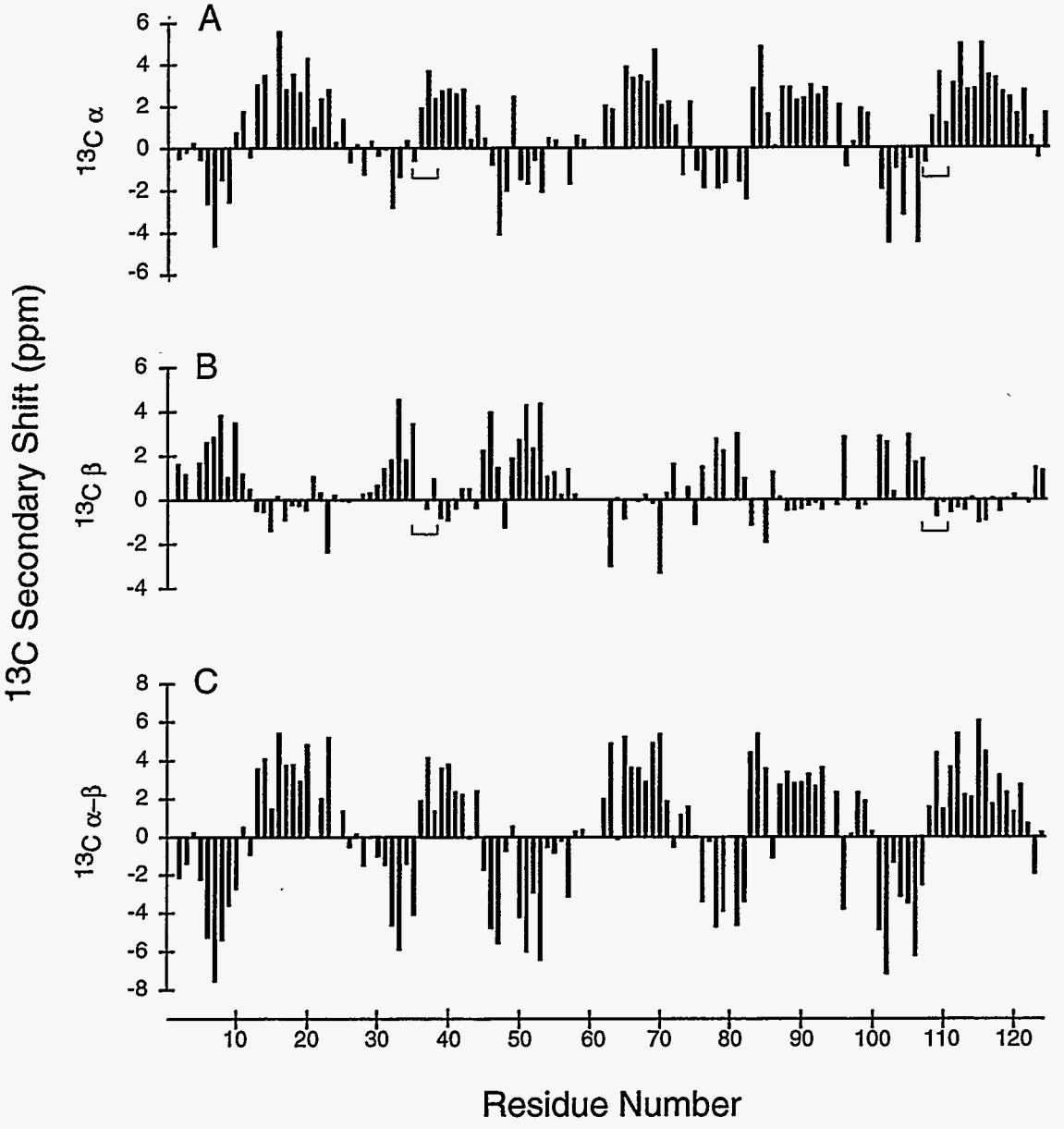


Figure 2.13: Observed ^{13}C secondary shifts of NTRC receiver domain, plotted as a function of residue number. Values reflect measured ^{13}C chemical shifts minus the amino acid-specific "random coil" ^{13}C chemical shift (Spera & Bax, 1991) in ppm, with positive and negative values indicating downfield and upfield secondary shifts, respectively. (A) $^{13}C_\alpha$ secondary shifts. (B) $^{13}C_\beta$ secondary shifts. (C) $^{13}C_\alpha$ secondary shifts – $^{13}C_\beta$ 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.

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.

NOE's characteristic of N-terminal capping interactions have also been identified (Lyu et al., 1993; Zhou et al., 1994). 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 2.14 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 2.9. The N-terminal residue of strand 3, D49, plays an unusual role

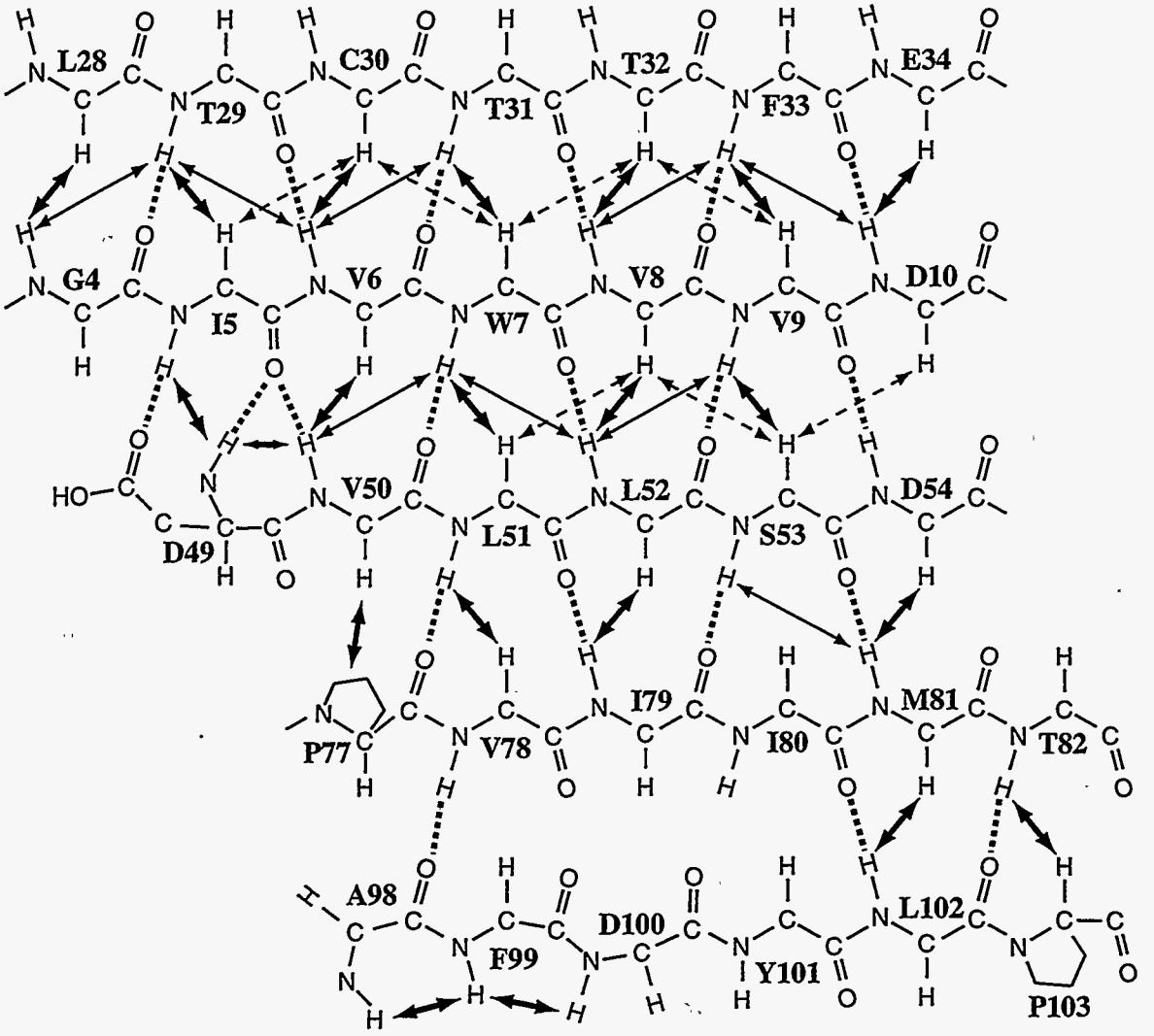


Figure 2.14: Cross-strand NOEs and hydrogen bonds observed for the β-sheet of NTRC receiver domain. Solid arrows correspond to NOEs observed in 3D ^{15}N NOESY-HMQC and dashed arrows correspond to NOEs observed in 2D ^{1}H-^{1}H NOESY. Dashed lines indicate hydrogen bonds included in structure calculations involving amide protons with reduced exchange rates.

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 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 Hα-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 2.2 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 82 restraints defining 41 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

Table 2.2 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.

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 average of 0.81 Å for the backbone atoms and 1.35 Å for non-hydrogen atoms. Figure 2.15 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 2.16. 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 D11, 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 D12 is not well-defined in the family of structures due, also, to a lack of NOEs for that residue.

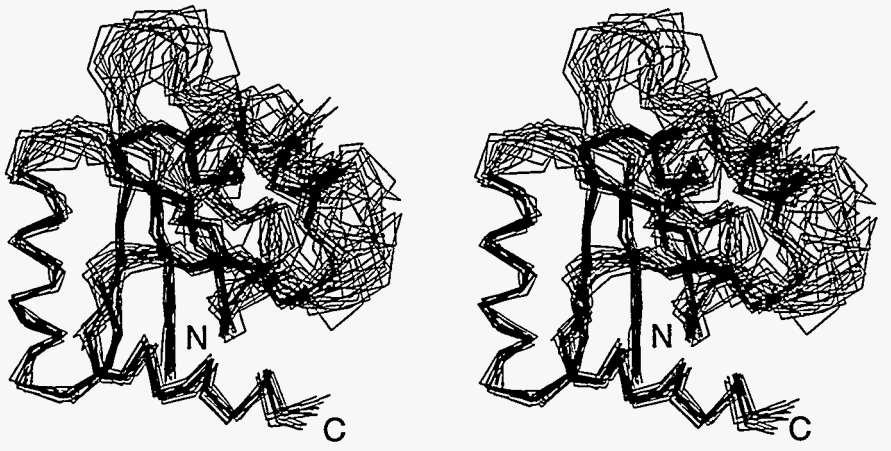


Figure 2.15: Stereoview of the family of 20 distance geometry-simulated annealing structures of NTRC receiver domain. Structures are superimposed on backbone atoms of the average structure, including residues 4-10, 14-44, 48-55, 65-73, 77-82, and 98-121.

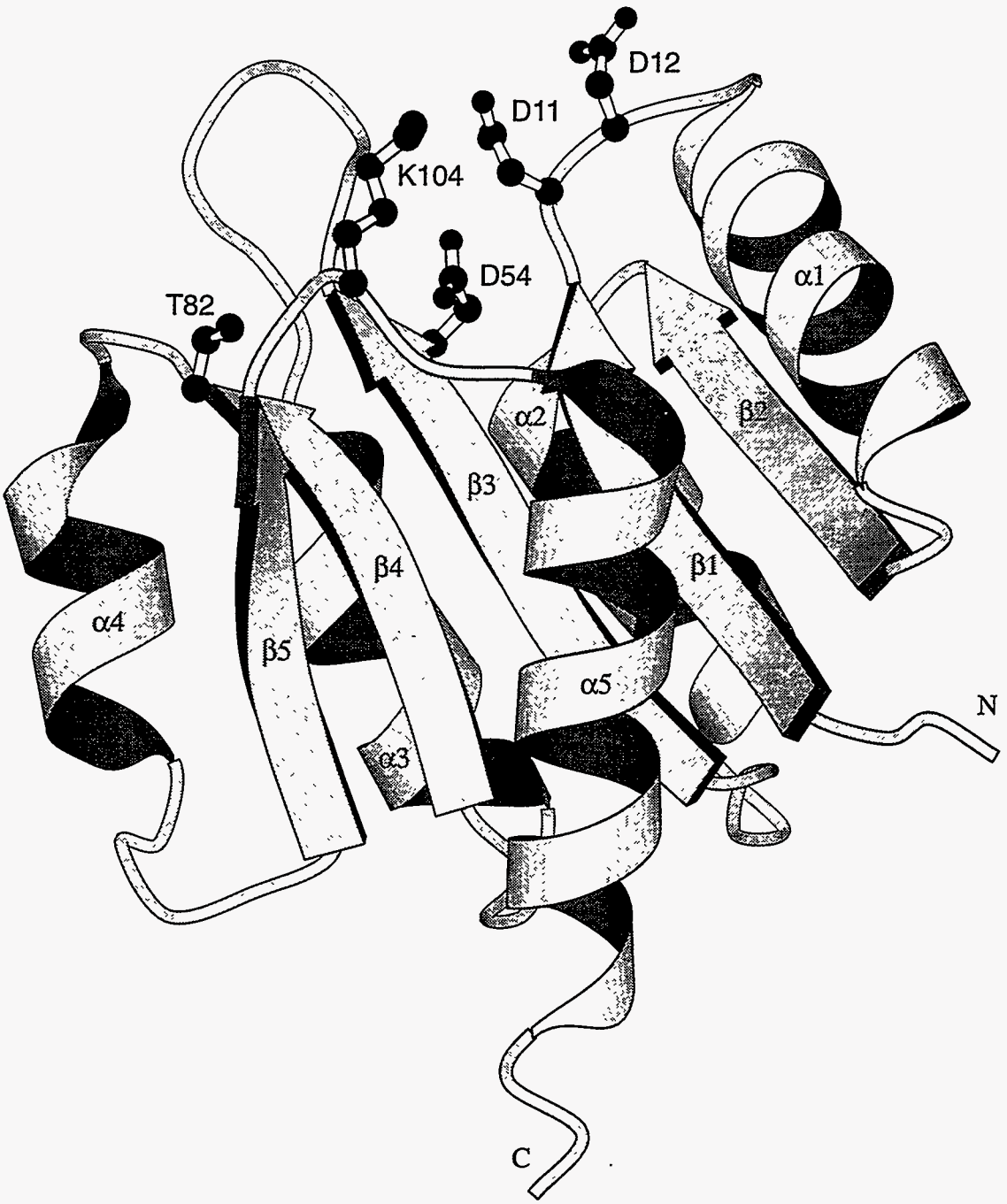


Figure 2.16: Ribbon diagram of the NTRC receiver domain. Secondary structure elements are individually labeled. Active site residues (D10, D11, D54, T82, K104) are labeled and shown in ball-and-stick representation.

NTRC is shown from different angles in Figures 2.15 and 2.16. This was done to highlight the structured regions in Figure 2.15 and the active site in Figure 2.16. The view of Figure 2.16 can be converted into the view in figure 2.15 by two 90° rotations. The first rotation is about a horizontal axis and moves the active site towards the viewer. The second rotation is about a vertical axis and moves the active site to the left of the page.

Discussion

Comparison of NTRC receiver domain with CheY

High-resolution structures of CheY, a homologous receiver domain protein involved in chemotaxis, 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 C_α 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 elements. Orientations of helices 1, 2, 3, and 5 relative to the sheet are generally similar in the NTRC receiver domain and CheY. The C_α 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 2.17A shows the superposition of the 20 DGSA structures of the NTRC receiver domain on the C_α trace of the crystal structure of CheY in the absence of Mg^{2+} (Brookhaven PDB file 3CHY). Average displacements from the corresponding CheY C_α coordinates of 8.1 to 9.5 Å are observed for the C_α atoms of residues L87 to A90 of NTRC. The RMS deviations of those four C_α 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 2.17B 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: D11, D12, D54, T82 and K104 (Moore et al., 1993; Volz, 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 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

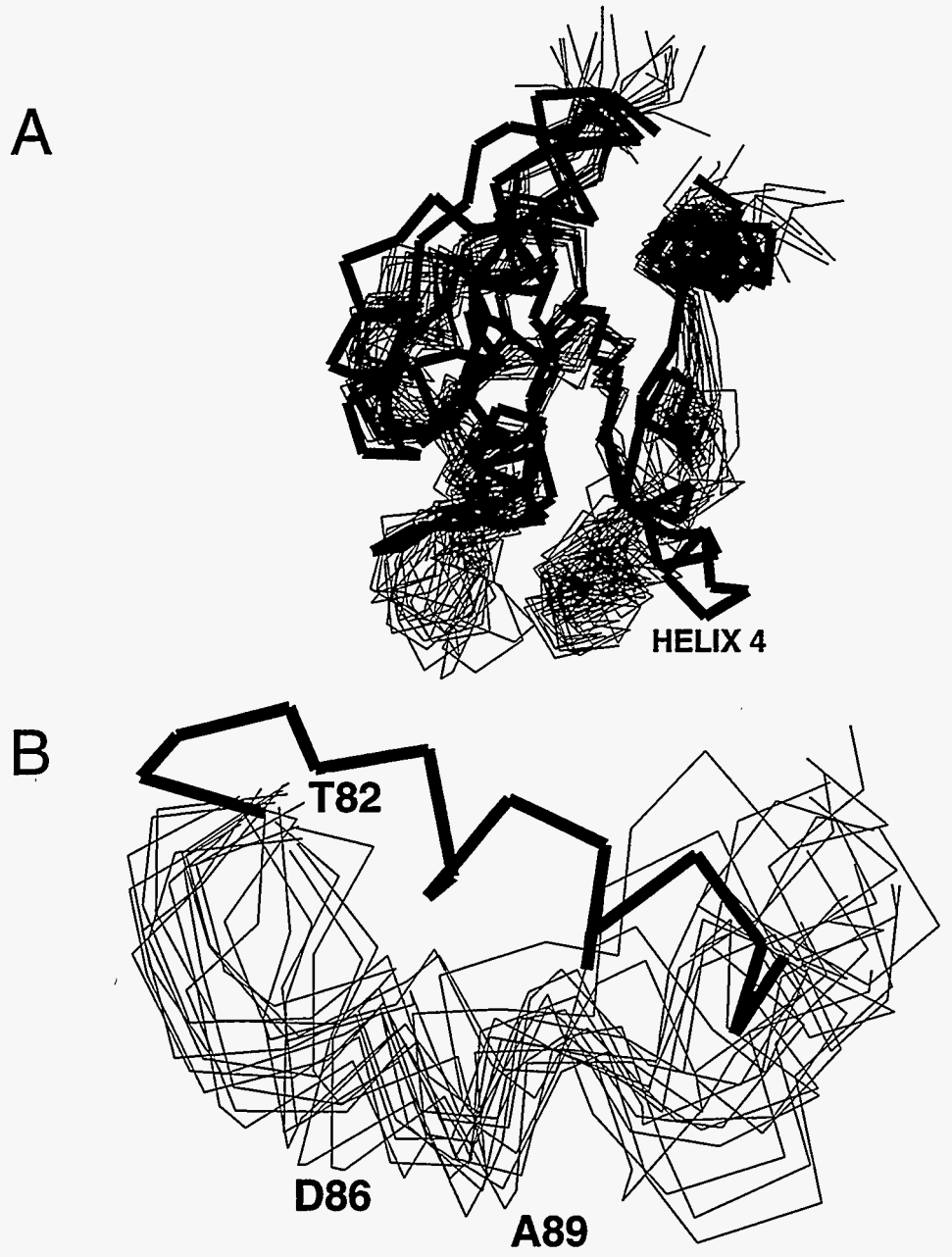


Figure 2.17: Superposition of 20 NTRC receiver domain structures on the high resolution crystal structure of CheY, demonstrating the difference in the position of helix 4 in the two proteins. The residues of all elements of secondary structure except for helix 4 were used for the alignment of structures. 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.

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 (Flashner et al., 1995). 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).

Summary

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.

Chapter 3 Refined Solution Structure and Backbone Dynamics of Unphosphorylated NTRC

Introduction

Overview

One of the most interesting aspects of the NTRC N-terminal domain project is the structural basis of regulation of this domain by phosphorylation. In order to investigate this question, structures in both the inactive unphosphorylated and active phosphorylated forms will need to be compared. To maximize the likelihood of detecting subtle structural differences between the forms, it was decided to pursue a higher resolution structure of the unphosphorylated form. Furthermore, a portion of the intermediate resolution structure presented in chapter 2, comprising helix 4 and strand 5, displayed a low NOE density. This can be indicative of unusual dynamic motions in secondary structure. Thus, analysis of backbone dynamics of the unphosphorylated form of NTRC were undertaken.

Refinement

The structure presented in chapter 2 shows very good RMSD's for the backbone atoms involved in secondary structure at 0.81 $\pm$ 0.06 Å (see Table 2.2). However, as is typical of intermediate resolution structures the sidechains are less well defined at 1.35 $\pm$ 0.11 Å, for all non - H atoms in regions of secondary structure. Note that these numbers were determined without helix 4 and strand 5 as these regions of the molecule appear to have a low NOE density that can be indicative of increased dynamic behavior.

The sidechains are relatively poorly defined because the intermediate resolution structure presented in chapter 2 lacks stereospecific assignments for prochiral groups. The most important stereoassignments for the determination of higher resolution structures are the methyl groups of leucine and valine (Guntert et al., 1989). Without the ability to distinguish between prochiral groups, distance constraints involving these groups are defined to a pseudo-atom that lies between the protons involved (Guntert et al., 1989).

This adds a large upper bound correction to the distance restraint which lowers the precision of the calculated structures.

Also, a truly high resolution structure will have 15-20 restraints per residue. The structure presented in chapter 2 has about 10 restraints per residue. 1620 peaks were picked from the 4D NOESY of which about 800 peaks were assigned in the 3-4 passes through the data that were performed to generate the intermediate structure. Thus, it seemed prudent to continue screening the other 800 peaks with better tools to try to extract more restraints.

Backbone Dynamics

The intermediate resolution structure of NTRC has a low NOE density for the region of the structure involving helix 4 and strand 5 which resulted in poor convergence in the final structures for those regions of the protein. It is not uncommon to see poor convergence for loops in solution structures, and these have often been shown to be regions of high mobility (Mandel et al., 1995). However, it is very unusual to find such high mobility in regions of secondary structure. It was decided to use backbone dynamics to probe whether this paucity of NOE's involving helix 4 and strand 5 could be correlated to fast time scale motions. This was carried out by the analysis of the relaxation of backbone amide nitrogens in uniformly ^{15}N labeled protein. The dynamic behavior of a number of proteins have been analyzed by this method (Palmer, 1993; Wagner, 1993).

In one of the most informative cases, this type of dynamic information resolved a long standing question about the central linker domain in calmodulin. A crystallographic study had suggested that this central linker is a continuous helix (Babu et al., 1988) while various small-angle X-ray scattering experiments in solution produced conflicting information about the flexibility of this region (Seaton et al., 1985; Heidorn & Trewhella, 1988; Matsushima et al., 1989). A study of the backbone dynamics of calmodulin showed that the central domain is indeed flexible and suggested that the role of the central linker

was to simply act as a tether between the two globular binding domains (Barbato et al., 1992).

Material and Methods

Expression and Purification of NTRC

The expression and purification scheme presented in this section is similar to that described in chapter 2. However, this scheme is superior in yield and purity.

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. In the case of the 10% labeled sample, the cells were grown in a 90% ^{12}C/ 10% ^{13}C glucose mix. Production of the NTRC receiver domain was induced with the addition of 1 mM isopropyl β-D-thiogalactopyranoside after the cell density had reached 0.8 absorbance units at 595 nm. The cells were grown for 7-8 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 sodium 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 Amicon ultrafiltration concentrator (Amicon). Final HPLC purification was performed on a 1.66 ml HQ/M ion exchange column on a BioCad Sprint FPLC system (PerSeptive BioSystems). Purity and identity of the protein were confirmed by mass spectroscopy, gel electrophoresis and NMR.

Sample Preparation

Concentrated protein solution was flow dialyzed against 10 mM 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 0.1 M HCl or NaOH. The concentration of the uniformly ^{13}C, ^{15}N-labeled sample was 2 mM. The concentrations of NMR samples are based on the weight of the lyophilized material after HPLC purification and dialysis against water and UV absorption at 280 nm (NTRC extinction coefficient - 14060 M^{-1} cm^{-1}) (Gill & von Hippel, 1989).

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 and all data were processed as to yield 512 x 512 matrices.

^{15}N spin-echo difference constant time HSQC (Vuister et al., 1993) and ^{13}CO spin-echo difference constant time HSQC (Grzesiek et al., 1993) experiments were collected with spectral widths of 6944 Hz for the ^1H dimension and 5000 Hz for the ^{13}C dimension. The experiments were centered at 43.16 ppm in ^{13}C and 4.78 ppm in ^1H.

^1H-^{13}C HSQC (Otting & Wüthrich, 1988) experiments were collected with spectral widths of 6944 Hz for the ^1H dimension and 5200 Hz for the ^{13}C dimension. The experiments were centered at 45.3 ppm in ^{13}C and 4.78 ppm in ^1H.

T_1, T_2 and NOE relaxation measurements (Skelton et al., 1993a) were taken with a spectral width of 6944 Hz in ^1H and 1861 Hz in ^{15}N. The spectra were centered at 4.78 ppm in ^1H and 119.1 ppm in ^{15}N. The recycle delay was set to 1.5 s for the T_1, T_2, and

NOE saturation experiments. The NOE reference experiment was taken with a 8.5 s recycle delay to allow the bulk water to undergo significant T_1 relaxation.

T_2 measurements were taken with relaxation delays of 7.84 ms, 39.2 ms (x2), 125.44 (x2), 156.8 ms, 309.76 ms (x2), and 501.74 ms. T_1 measurements were taken at 26 ms (x2), 154.2 ms, 324.51 ms (x2), 537.5 ms (x2), 821.6 ms, 1390.3 ms (x2).

Stereoassignments

Stereoassignments of valine methyls were obtained from a combination of the ^{15}N spin-echo difference constant time HSQC (J_{NC}) (Vuister et al., 1993) and ^{13}CO spin-echo difference constant time HSQC (J_{CC}) (Grzesiek et al., 1993) experiments. The J_{CC} experiment gives the 3 bond coupling between the carbonyl and γ carbon of a residue. The J_{NC} experiment gives the 3 bond coupling between the amide nitrogen and γ carbon of a residue. These experiments are particularly effective for isoleucine, valine and threonine because these residue contain γ methyls which appear in a well resolved portion of the spectrum and are quite intense. Measurement of both coupling constants can determine the relevant χ_1 angle and, in the case of valine, provide stereospecific assignment of the γ methyls.

The J_{CC} and J_{NC} experiments are members of the quantitative J coupling class of experiment. In these experiments, the coupling constants are determined from the magnetization loss due to dephasing caused by unresolved J couplings. Two constant time experiments are taken in an interleaved fashion in this scheme. In the reference experiment, the coupling of interest (J_{CC} or J_{NC}) is suppressed by a 180° pulse on the appropriate channel during the carbon constant-time evolution period. In the second experiment, this 180° pulse is shifted to a position that allows the coupling of interest to be active for the carbon constant-time evolution period which causes attenuation of the magnetization. The attenuation is governed by $\cos\left(\pi J_{NC/CC} T\right)$. The value of the coupling constant is calculated from a ratio involving the signal from the reference experiment, S_a, and the signal from the attenuated experiment, S_b:

$$\frac{(S_a - S_b)}{S_a} = 1 - \cos(\pi J_{NC/CC} T) = 2\sin^2(\pi J_{NC/CC} T) \qquad (3.1)$$

Stereospecific assignments for leucine methyls and valine methyls were obtained by biosynthetically directed fractional labeling (Neri et al., 1989). In this method, a sample is prepared with a 10% ^{13}C/90% ^{12}C glucose mix. The *pro-R* methyl and the adjacent methine group are made from the same pyruvate during biosynthesis and, thus, both will be labeled with ^{13}C in 10% of the molecules. The *pro-S* methyl and the adjacent methine originate from different pyruvate molecules and, as such, are only expected to both be labelled in 1% of the molecules. Therefore, these methyl groups can be distinguished by an HSQC without constant-time broadband homonuclear decoupling during the ^{13}C evolution period. The *pro-R* methyl group will be split by about 35 Hz due to the carbon-carbon coupling between the methine and methyl while the *pro-S* methyl appears as a single peak.

Dynamic Analysis

The T_1 and T_2 data for each residue were calculated using a non-linear least squares fitting algorithm from Mikael Akke in Art Palmer's laboratory at Columbia University. The heteronuclear NOE parameter was determined by taking the ratio of the intensity of the saturation experiment to the reference experiment for each amide.

Uncertainties in T_1 and T_2 measurements were determined from the duplicate points. Uncertainties in the NOE data were derived from the standard deviation for noise in the spectrum determined from a region of the spectrum without signal (Skelton et al., 1993a).

The relaxation of a ^{15}N amide nucleus is dominated by chemical shift anisotropy and dipolar coupling at high field. Three easily obtained parameters of this relaxation are T_1, T_2 and the heteronuclear NOE (Skelton et al., 1993b). These relaxation parameters can

be described for the amide ^{15}N in terms of the spectral density at relevant frequencies (see chapter 1) (Abragam, 1961):

$$\frac{1}{T_1} = \left(\frac{d^2}{4}\right)\left[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)\right] + c^2 J(\omega_N) \quad (3.2)$$

$$\frac{1}{T_2} = \left(\frac{d^2}{8}\right)\left[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)\right] + \left(c^2/6\right)\left[4J(0) + 3J(\omega_N)\right] + R_{ex} \quad (3.3)$$

$$NOE = 1 + \left(\frac{d^2 T_1}{4}\right)\left(\frac{\gamma_H}{\gamma_N}\right)\left[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)\right] \quad (3.4)$$

where

$$d = \left(\frac{\mu_o h \gamma_H \gamma_N}{8\pi^2}\right)\left\langle r_{NH}^{-3}\right\rangle \quad (3.5)$$

$$c = \omega_N(\sigma_\parallel - \sigma_\perp)/\sqrt{3} \quad (3.6)$$

where μ_o is the permeability of free space, h is Planck's constant, γ_H is the gyromagnetic ratio of hydrogen, γ_N is the gyromagnetic ratio of nitrogen, r_{NH} is the amide bond length (1.02 Å), ω_H and ω_N are the Larmor frequencies of ^1H and ^{15}N,

$\sigma_{\|}$ and $\sigma_{\perp}$ are the parallel and perpendicular components of the chemical shift anisotropy tensor. R_{ex} is used to account for chemical exchange in (3.3).

Information about the internal dynamics of NTRC was obtained by the model-free formalism of Lipari and Szabo (Lipari & Szabo, 1982b; Lipari & Szabo, 1982a; Clore et al., 1990). In this analysis the spectral density function, $J(\omega)$, can be modeled as:

$$J(\omega)=\frac{2}{5}\left[\frac{S^2\tau_m}{1+(\omega\tau_m)^2}+\frac{(1-S_f^2)\tau_f'}{1+(\omega\tau_f')^2}+\frac{(S_f^2-S^2)\tau_s'}{1+(\omega\tau_s')^2}\right] \quad (3.7)$$

where

$$\tau_f'=\tau_f\tau_m/(\tau_f+\tau_m) \quad (3.8)$$

$$\tau_s'=\tau_s\tau_m/(\tau_s+\tau_m) \quad (3.9)$$

where τ_m is the overall correlation time, τ_f is the effective correlation time for fast time scale internal motions (less than 150 ps), τ_s is the effective correlation for slow time scale internal motions ($\tau_m > \tau_s > \tau_f$), S is the generalized order parameter $\left(S^2 = S_f^2 S_s^2\right)$, S_f is order parameter for fast internal motions, and S_s is the order parameter for slow internal motions.

There are actually six possible models of the spectral density as a function of an order parameter. The model presented in (3.7) must fit six parameters (τ_m, τ_s, τ_f, S^2, S_f^2, and R_{ex}) and will henceforth be referred to as model 6. There are five simpler models (models 1-5). Model 1 (S^2, τ_m) is obtained by assuming that their are no significant motions on the slow time scale and the motions on the fast time scale are very fast (<

20 ps). Model 2 (S^2, τ_m, Rex) is model 1 with chemical exchange. Model 3 (S^2, τ_m, τ_f) is obtained by assuming that their are no significant motions on the slow time scale. Model 4 (S^2, τ_m, τ_f, Rex) is model 3 with chemical exchange. Model 5 (S_s^2, S_f^2, τ_m, τ_s) assumes significant motion on both the fast and slow time scales. Model 6 assumes motion on both fast and slow time scales and exchange term.

One of the trickiest questions for model-free analysis is the determination of the appropriate model. One would like to choose the most parsimonious model that fits the relaxation data (T_1, T_2 and NOE) for each amide. Models were chosen on the basis of the T_1/T_2 and NOE values (Kay et al., 1989; Clore et al., 1990; Redfield et al., 1992). A total of 45 residues with a heteronuclear NOE of greater than 0.75 and a T_1/T_2 ratio within one standard deviation from the average, 7.20 ± 0.7, were analyzed with model 1. A total of 10 residues with a heteronuclear NOE greater than 0.75 and a T_1/T_2 ratio greater than one standard deviation from the average, 7.20 ± 0.7, were analyzed with model 2 which adds an exchange term to model 1. A total of 11 residues with a heteronuclear NOE below 0.75 and a T_1/T_2 ratio within one standard deviation from the average, 7.20 ± 0.7, were analyzed with model 3. A total of 13 residues with a heteronuclear NOE below 0.75 and a T_1/T_2 ratio greater than one standard deviation from the average, 7.20 ± 0.7, were analyzed with model 4 which adds an exchange term to model 3. A total of 3 residues with a heteronuclear NOE below 0.75 and a T_1/T_2 ratio less than one standard deviation from the average, 7.20 ± 0.7, were analyzed with model 5 which assumes significant motion on two time scales.

Model-free parameters were fit against the experimental data (3.2-3.4) using the program DNMR version 3.1(Orekhov et al., 1995). An initial value of the total correlation time, τ_m, was estimated from the 10% trimmed mean of T_1/T_2 for the amides to be 9.5 ± 0.25 (Kay et al., 1989).

Results

Stereospecific Assignments

The coupling constants calculated for the J_{CC} and J_{NC} experiments are shown in Table 3.1. They are calculated as described in the Material and Methods. χ_1 angles were determined by treating the coupling constants as either large (>1.5 for the J_{NC} and >2.0 for the J_{CC}) or small (<1.0 for the J_{NC} and <1.5 for the J_{CC}) (Grzesiek et al., 1993; Vuister et al., 1993). A large coupling constant indicates that the atoms involved are in a *anti* rotamer state with respect to each other while a small coupling constant indicates a *gauche* rotamer state. If neither coupling constant is large, the presumption is that there is rotamer averaging between the two states.

The data shown in Table 3.1 contain a few of ambiguous situations where both the J_{NC} and J_{CC} coupling constants are large (I15, I119). These situations are hard to interpret in the framework presented above. It is worth noting that one possible source of this ambiguity could lie in the J_{CC} experiment. This experiment was published with the use of pulsed field gradients (Grzesiek et al., 1993). The laboratory did not have the capability to implement this sort of experiment at the time the J_{CC} was run. Therefore, the experiment was implemented with phase cycling replacing the pulsed field gradients for artifact suppression (Bax & Pochapsky, 1992). This may have lead to imperfect artifact suppression thus skewing the values of the calculated coupling constants.

For the most part, however, the coupling constants allowed determination of the relevant χ_1 angles. A number of valine residues displayed behavior typical of rotamer averaging. It is interesting to note that these residues, V6, V18 and V39, fall into the well-structured portion of the protein. However, it has been shown in at least one case that rotamer averaging of valine sidechains can be correlated with low order parameters for the methyl groups (Kay et al., 1996). In the case of V78, V91 and V115 the coupling constants allowed determination of stereospecific assignments. These data are shown in Table 3.2.

Table 3.2 also contains the stereospecific assignments obtained from biosynthetically directed fractional labeling experiment. This experiment provided

Table 3.1: Coupling Constants Extracted from the J_{CC} and J_{NC} Experiments and Corresponding χ_1 angles.

Residue[a]	J_{NC} (hz)[b]	J_{CC} (hz)[b]	χ_1[c]
I15	2.00	1.76	-
I69	-	1.50	-60°
I79	0.75	2.00	-60°
I108	-	1.52	-60°
I119	1.74	1.27	-
T29	1.16	1.58	rotamer ave.
T31	0.98	2.14	+60°
T32	2.14	-	-60°
T82	-	0.87	-60°
V6(γ_1)	-	2.20	rotamer ave.
V6(γ_2)	0.72	1.53	rotamer ave.
V18(γ_1)	-	0.40	rotamer ave.
V18(γ_2)	1.07	0.55	rotamer ave.
V39(γ_1)	-	0.11	rotamer ave.
V39(γ_2)	0.95	1.73	rotamer ave.
V78(γ_1)	-	3.3	180°
V78(γ_2)	2.3	1.6	180°
V91(γ_1)	1.79	1.13	180°
V91(γ_2)	-	3.13	180°
V112(γ_1)	-	3.8	180°?
V112(γ_2)	-	-	180°?
V115(γ_1)	2.01	0.7	180°
V115(γ_2)	-	2.53	180°

[a] The valine designations are not stereospecific. They are the same designations given in table 2.1 and are used only as a label for distinguishing between the prochiral methyls.

[b] Errors are on the order of 0.1 Hz.

[c] The appelation Rotamer Ave. indicates residues which sample both the *anti* and *gauche* rotamer states.

stereoassignments for 7 out of 12 leucines for which both methyls were previously assigned and for 2 valines. One of these valines, V115, was stereospecifically assigned both by this method and by the couplings extracted from the J_{CC} and J_{NC} experiments. These methods agreed on the stereospecific assignment for this residue which provides a good internal control. The data from the biosynthetically directed fractional labeling experiment are somewhat tentative because the experiment was taken at a rather low resolution which resulted in the 35 Hz splittings used for stereospecific assignment to be barely resolved. The experiment should be repeated with higher resolution in the ^{13}C dimension to confirm the data in Table 3.2.

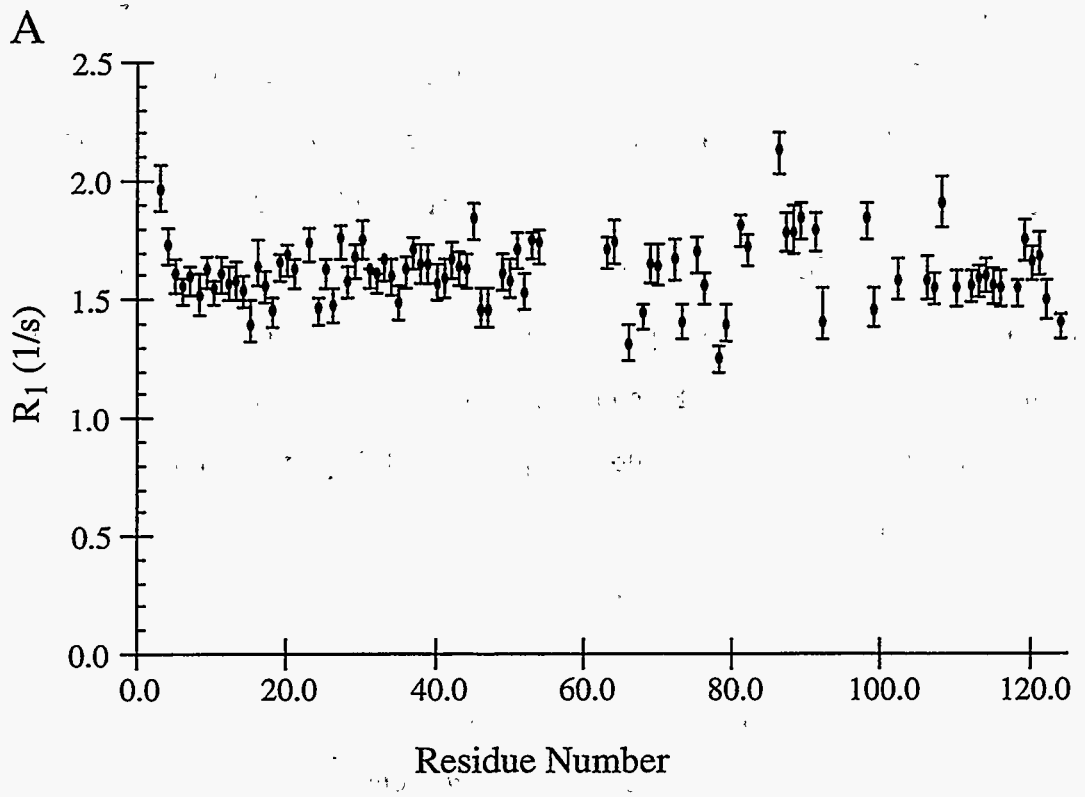
Further Restraints from the 4D $^{13}C/^{13}C$ NOESY

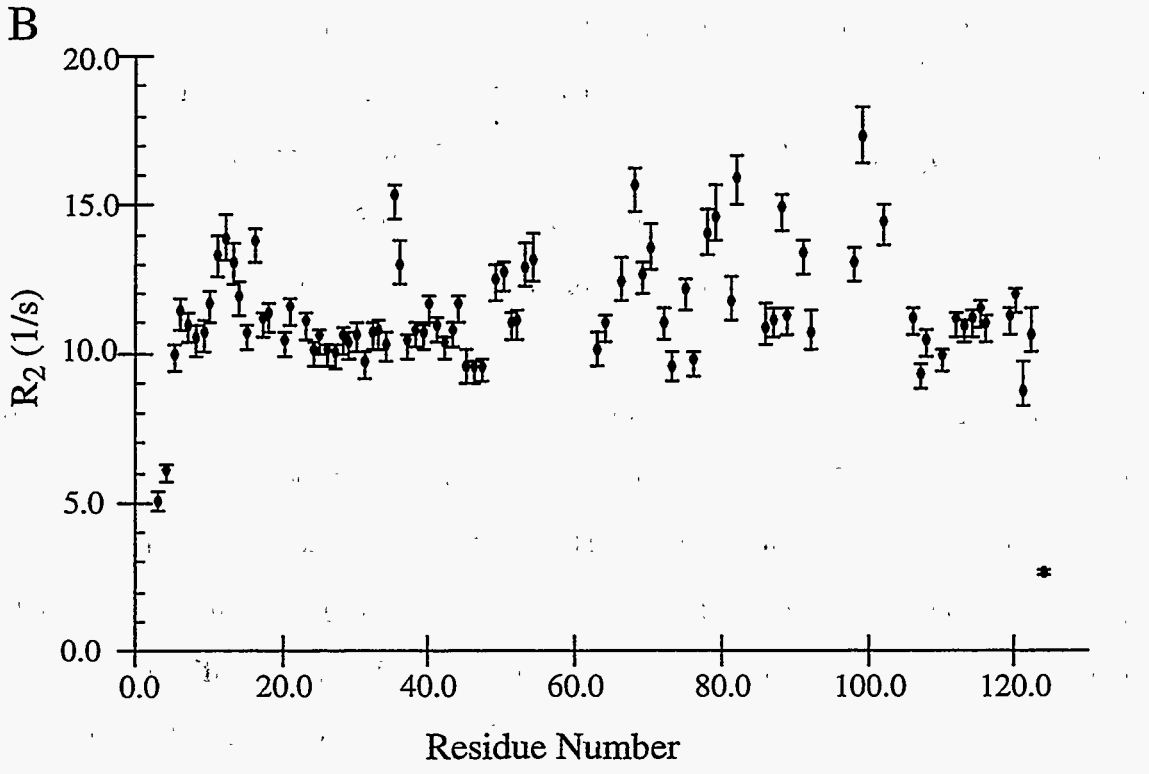
About 800 of the 1616 peaks in the 4D 13C/13C NOESY (see chapter 2) were assigned for the intermediate structure. Further screening of this experiment yielded a number of additional restraints. This process was carried out with a C program written by Jeff Pelton in the Wemmer laboratory. This program is an enhanced version of the program used in chapter 2, and not only matches peaks in the 4D NOESY with assignments, but also reads in a *pdb* (protein data bank) file and prints out the distance for each possible assignment. This greatly facilitates the refinement process.

A total of 83 candidate restraints have been identified by this method. These restraints have not yet been tested in rounds of structure calculations.

Backbone Dynamics

Figure 3.1A shows the ^{15}N longitudinal relaxation rates, R_1 ($1/T_1$), as a function of residue number for NTRC. These values are fairly uniform, although there is increased variation in the helix 3 - strand 5 region. The trimmed weighted mean of R_1 is 1.65 s^{-1}. Figure 3.1B shows the ^{15}N transverse relaxation rates, R_2 ($1/T_2$), as a function of residue number for NTRC. These values show considerable variability. Specifically, the loops between strand 1 and helix 1 as well as the region from helix 3 to strand 5 show increased R_2's. The trimmed weighted mean of R_2 is 11.25 s^{-1}.


A
2.5
2.0
1.5
1.0
0.5
0.0
R1 (1/s)
0.0
20.0
40.0
60.0
80.0
100.0
120.0
Residue Number


B
20.0
15.0
10.0
5.0
0.0
R2 (1/s)
0.0
20.0
40.0
60.0
80.0
100.0
120.0
Residue Number

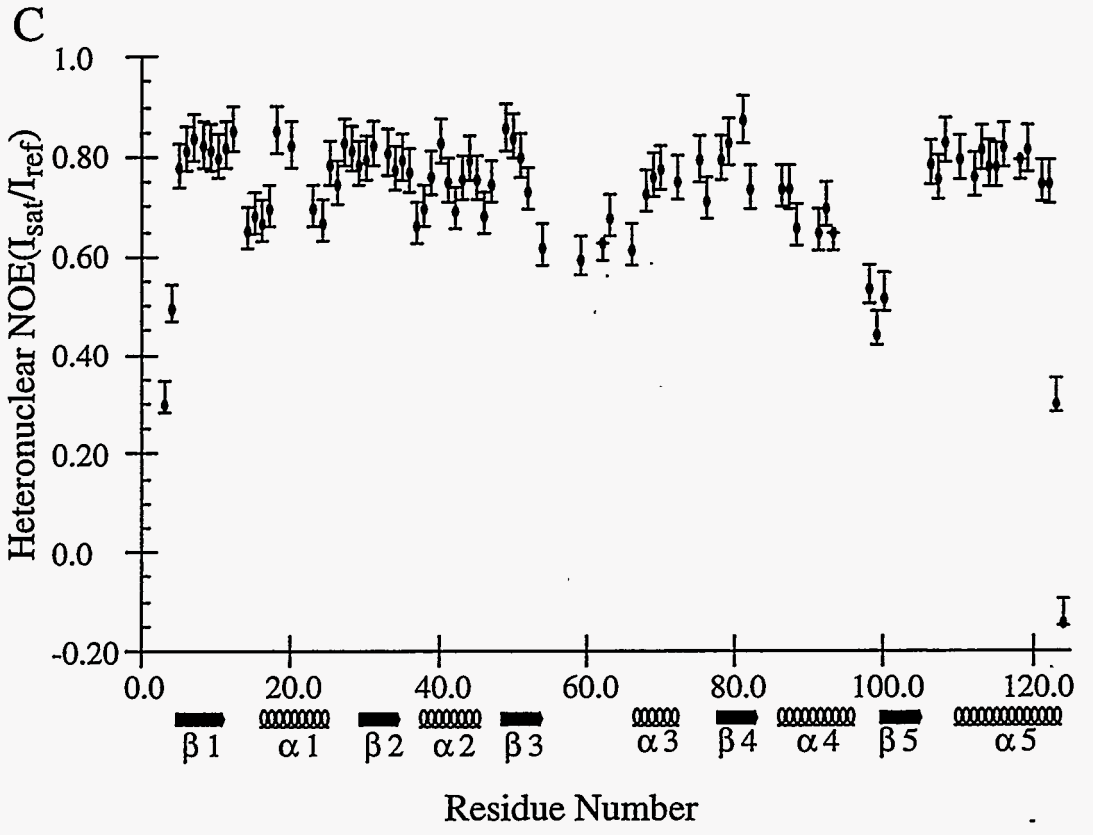


Figure 3.1: ^{15}N backbone relaxation parameters for the N-terminal domain of NTRC. Shown are (A) R_1 ($1/T_1$), (B) R_2 ($1/T_2$) and (C) Heteronuclear NOE. See the text for experimental details. Note that data could not be obtained for about 25 residues either due to severe overlap or the complete lack of a resonace for the residue. The missing peaks are probably due to amide exchange. The 10% trim weighted mean results were 1.65 s^{-1} for R_1, 11.25 s^{-1} for R_2, and 0.81 for the NOE.

Figure 3.1C shows the heteronuclear NOE values as a function of residue number for NTRC. The trimmed weighted mean of the heteronuclear NOE was 0.81. This experiment is often sufficient to detect regions of increased mobility if the motions are on the nanosecond-picosecond time scale. In particular, the loops between strand 3 and helix 3 (D54, the site of phosphorylation lies at the beginning of this loop) and strand 1 and helix 1 show lowered heteronuclear NOE's. Helix 4 and, most dramatically, part of strand 5 also show lowered NOE's. It is interesting to note that helix 3 and strand 4 do not show lowered NOE's which is in contrast to the behavior of these residues in the R_2 experiments. This could be indicative of slow timescale motions.

The heteronuclear NOE is very sensitive to amide exchange in the reference experiment. If protons from the water are not allowed to completely relax back to equilibrium during the recycle delay, these proton will exchange with the amide protons and cause a heteronuclear NOE effect. Thus, one will consistently underestimate the intensity of the reference experiment which will lead to a consistent overestimation of the heteronuclear NOE relaxation parameter. This, in turn, leads to an interpretation that the structure has less flexibility than is actually the case. The initial heteronuclear NOE experiments taken on NTRC suffered greatly from this problem. Many of the values were over the theoretical maximum (Mandel et al., 1995) and this caused a great deal of difficulty in fitting the data with the Lipari and Szabo formalism. This problem was circumvented by using an 8 s recycle delay for the reference experiment which allows most of the water magnetization to relax to equilibrium.

The results of the Lipari and Szabo model-free analysis are shown in Figure 3.2. A large number of residues (see Material and Methods for the specific residues involved) needed to be modeled with a chemical exchange term (models 2 and 4). Model 4 was invoked because these residues had a lowered heteronuclear NOE coupled with longer R_2's (i.e. faster relaxation). A lowered heteronuclear NOE indicates fast time scale motions that

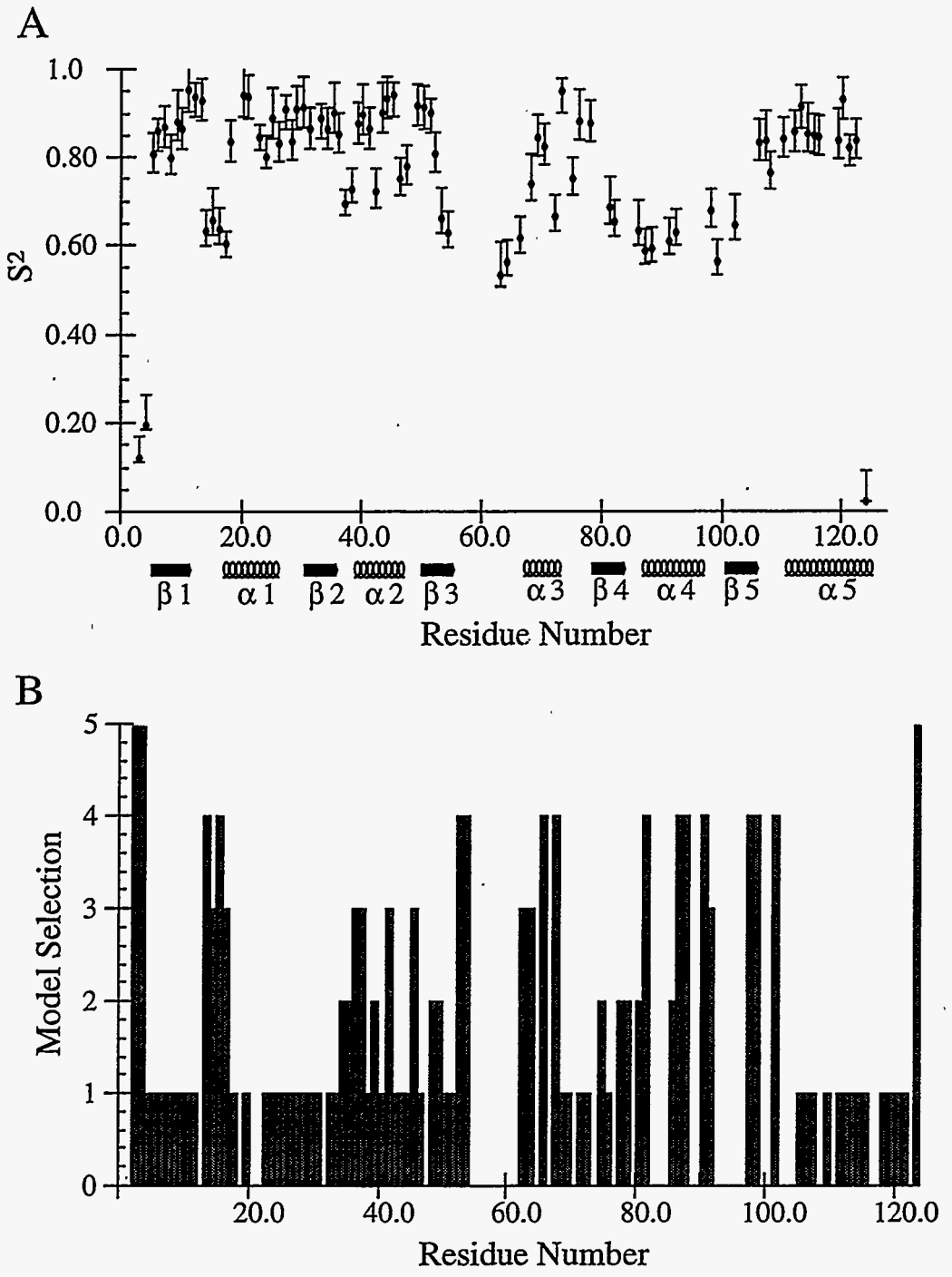


Figure 3.2: Model free parameters for the N-terminal domain of NTRC. Shown are (A) the order parameters for the backbone amides as a function of residue number and (B) The model of the spectral density used for analysis as a function of residue number. The models used (model 1-5) are described in the text. Note that a model selection of 0 indicates that complete relaxation data (R_1, R_2 and heteronuclear NOE) were not obtained for that residue.

should lower the internal correlation time and slow the rate of transverse relaxation. The fact that the rate of relaxation increased in these residues can be explained by the presence of chemical exchange processes. Similarly, those residues analyzed with model 2 had average heteronuclear NOEs, but increased R_2's which again indicated the need for an exchange term in (3.3) The model of the spectral density used for analysis as a function of residue number is shown in Figure 3.2B.

Figure 3.2a shows the generalized order parameters, S^2, determined for each residue. Order parameters vary between 0 and 1 and are a measure of the internal spatial flexibility of the N-H bond vector. An order parameter of 0 corresponds to completely isotropic motion of the bond vector while an order parameter of 1 corresponds to a completely rigid bond vector. The order parameters generally fall around 0.85 indicating a rigidly structured backbone. The loops between strand 1 and helix 1 and strand 3 and helix 3 as well as helix 4 and strand 5 and the loop following strand 5 show order parameters of about 0.625 indicating some internal motion in these regions. The N-terminal and C-terminal residues have extremely low order parameters which, unsurprisingly, indicates that theses residues are undergoing extensive internal motion.

Discussion

Progress Towards a High Resolution Solution Structure

The stereospecific assignments for 7 leucine and 4 valine residues have been obtained. This will allow removal of the psuedo-atom corrections for restraints involving these atoms which should improve the precision of the calculated structures. Since valine and leucine residues are usually buried, this will hopefully facilitate determination of a more precise set of structures for the core residues.

The new restraints that have been extracted from assignments of the 4D $^{13}C/13C$ NOESY should also improve the overall precision and accuracy of the NTRC structure. It must be noted that many of these restraints may not prove to be correct. This can easily be determined by evaluating the residual energies and convergence of structures calculated

with the new data. A winnowing process involving rounds of structure calculations will be used to identify the correct restraints.

The Dynamics of NTRC

The results of the backbone dynamics presented in this chapter confirm that helix 4 and strand 5 are dynamically active. This would explain the low NOE density and poor convergence for these regions in the structure presented in chapter 2. It is unusual to see such low order parameters for a large region of secondary structure. One possibility is that this region of the molecule requires flexibility in order to undergo a conformational change upon phosphorylation and conversion to the active form of the protein. It is interesting to note that this region of the molecule has been implicated in conformational change upon activation (see chapter 5). Alternatively, this region of the molecule could normally be packed against the central domain of NTRC and making important contacts for its stability. When the N-terminal domain of NTRC is expressed alone these contacts would be lost which might lead to a "looser" structure. Of course, these explanations are not mutually exclusive.

There is also interesting dynamic behavior in the loops and turns between strand 1 and helix 1 (~ 5 residues), strand 3 and helix 3 (~ 12 residues), strand 4 and helix 4 (~ 3 residues, and strand 5 and helix 5 (~ 5 residues) . All of these loops lie at the C-terminal end of the parallel β-sheet structure and contain active site residues. The loop following strand 1 contains D11 and D12 which coordinate the Mg^{2+} binding (note that D11 and D12 are fairly rigid). The loop following strand 3 contains D54 which is the site of phosphorylation. The loop following strand 4 contains T82 - the possible function of which is discussed in chapter 5. Finally, the loop following strand 5 contains K104 which, by analogy to CheY, is involved in Mg^{2+} binding. Thus, it appears that the active site of NTRC displays dynamic behavior.

Furthermore, many of these residues in regions with low order parameters were analyzed using models containing exchange terms. This can be an indication of

microsecond-millisecond time scale motion. However, to truly quantify the exchange contribution to relaxation, relaxation experiments should be taken at different magnetic field strengths. It is not possible to obtain such data at the present time for NTRC because the requisite field strengths are not locally available. Thus, exchange parameters must be interpreted conservatively.

Summary

Stereospecific assignments for a significant number of valine and leucine residues have been obtained. A number of potential distance restraints have been extracted by further analysis of the 4D $^{13}C/^{13}C$ NOESY. These data will allow calculation of a high resolution solution structure of NTRC. The dynamic behavior of the backbone of NTRC has been determined from amide relaxation experiments. These data explain the poor convergence of helix 4 and strand 5 in the intermediate structure of NTRC. Furthermore, the dynamic analysis suggest that the active site of NTRC undergoes significant motion in the unphosphorylated form.

Chapter 4 Structural Studies of the Mg^{+} bound and Phosphorylated forms of NTRC

Introduction

Overview

This chapter presents the progress towards the determination of the structure of the phosphorylated form of NTRC. This has proven difficult to obtain due to the short lifetime of the phosphorylated form. Recently, conditions have been obtained that appear promising for structural analysis. Also discussed are the structural changes upon binding of Mg^{2+} to the N-terminal domain of NTRC.

Regulation of Proteins via Phosphorylation

One of the most common mechanisms of post-translational regulation of protein activity by phosphorylation of specific sidechains in the protein. Control of protein activity by phosphorylation is found in a diverse array of processes such as the cell cycle, transcription, translation, metabolic pathways, muscle contraction, memory, membrane transport, DNA replication, and signal transduction (Hunter & Sefton, 1991; Hardie, 1993). However, there is very little structural information about the structural basis of regulation by phosphorylation (Johnson, 1994). The only cases for which the 3 dimensional structures of both the phosphorylated and unphosphorylated forms of a protein have been determined are rabbit muscle glycogen phosphorylase (Sprang et al., 1988; Barford & Johnson, 1989; Johnson & Barford, 1993) and *E. coli* isocitrate dehydrogenase (Dean & Koshland, 1990; Hurley et al., 1990).

In the case of isocitrate dehydrogenase, the phosphorylation of a particular serine in the binding pocket sterically blocks the binding of citrate (Dean & D.E., 1990; Hurley et al., 1990). There is electrostatic repulsion between the phosphate and the carboxyl groups of the substrate. There are small local rearrangements near the site of phosphorylation, but no detectable changes distally.

In the case of rabbit muscle glycogen phosphorylase, the phosphorylation works by allosterism to stimulate the enzymatic activity (Sprang et al., 1988; Barford & Johnson, 1989; Johnson & Barford, 1993). The enzyme is a dimer in the inactive state, but aggregates to a tetramer upon phosphorylation. Phosphorylation at a serine about 30Å from the catalytic site causes large tertiary and quaternary conformational changes in the homodimer that lead to activation. Specifically, the 20 N-terminal residues go from a disordered state to a distorted 3_{10}-helix upon phosphorylation at serine 14. These N-terminal residues rotate about 120° and make contacts at the intersubunit surface. In doing so, they displace the five C-terminal residues and cause them to become disordered. Thus, upon phosphorylation, the N-terminal residues become ordered at the expense of disordering the C-terminal residues. The interdigitation of the N-terminal residues at the dimer interface causes large tertiary rearrangements that lead to a 10° rotation of the dimer subunits with respect to each other. This creates a new protein/protein interface and drives the formation of a tetramer. The phosphorylation appears to exert its effect through electrostatic interactions. The site of phosphorylation (*) is in a stretch of positively charged residues (RKQIS*VR) that are near a patch of acidic residues on the protein surface. Upon phosphorylation these N-terminal residues are moved from their previous site by electrostatic repulsion.

Phosphorylation in CheY

There have been several studies of CheY in the phosphorylated form. CheY was labeled on its six phenylalanines by incorporation of 4-fluorophenylalanine allowing analysis of ^{19}F chemical shift changes in the Mg^{2+} bound and phosphorylated states (Bourret et al., 1993; Drake, 1993). Binding of Mg^{2+} caused only local changes in the active site. Upon phosphorylation, long range conformational changes were observed to extend from the active site to phenylalanines on the other side of the protein. Unfortunately, this method of monitoring structural changes via chemical shift differences cannot describe the details of the conformational change (see chapter 5). The lifetimes for

the phosphorylated state obtained in this study were not sufficient for further structural analysis.

More recently, the backbone amide resonances of phosphorylated CheY were assigned by NMR (Lowry et al., 1994). This allowed comparison of chemical shift differences between the unphosphorylated and phosphorylated states at every residue except the prolines in the backbone. Large chemical shift differences, indicating conformational change, were observed in the active site, the end of helix 3, strand 4, helix 4, strand 5, and the beginning of helix 5 with smaller changes observed throughout the rest of the protein. Again, the very short half-life of the phosphorylated form of CheY, about 30 seconds, renders complete structural analysis difficult.

Small Molecule Donors

Bacteria that are lacking NTRB are able to activate transcription from *glnAp2* (Reitzer & Magasanik, 1985) which indicates that NTRC is capable of taking a phosphate from other donors. There are a large number of two-component histidine kinases in the cell that could also potentially act as donors for NTRC. It has been shown that phosphotransfer between non-cognate two-component systems is possible and in some cases it has been proposed to be physiologically relevant (Wanner, 1992).

It has also been shown that CheY and NTRC can phosphorylate themselves from small molecule donors such as acetyl phosphate, carbamyl phosphate and phosphoramidate, but not ATP or phosphoenolpyruvate (Feng et al., 1992; Lukat et al., 1992). Acetyl phosphate, in particular, may have some role in the regulation of the NTRC system as there are large pools of acetyl phosphate in bacterial cells. Phosphorylation of the N-terminal domain activates the ATPase in the central domain of full length NTRC and, in turn, activates transcription from the *glnAp2* promoter. The ability to use small molecules as phosphate donors suggests that receiver domains, such as NTRC and CheY, should be viewed as the active catalysts of their own phosphorylation.

Phosphorylation in NTRC

The phosphorylation site of NTRC has been mapped to aspartic acid 54 (Sanders et al., 1992). The phosphorylated state is intrinsically unstable in the native protein with a $t_{1/2}$ of about 4 minutes (Keener & Kustu, 1988). This has greatly limited the feasibility of structural studies on the phosphorylated form of the protein.

Magnesium Binding in Receiver Domains

There are two conflicting studies of CheY with Mg^{2+} bound by X-ray crystallography (Stock et al., 1993; Bellsolell et al., 1994). Bellsolell et al. observed large changes in the binding site and rearrangement and unwinding of the top of helix 4. Stock et al. did not observe this change and only detected a small rearrangement of the active site. The Stock et al. structure is more consistent with other investigations of the Mg^{2+} form of CheY (Drake et al., 1993;Bourret et al, 1993;Lowry et al., 1994). These NMR investigations suggested that there are not large conformational changes associated with Mg^{2+} binding.

Oligomerization in NTRC

There has been some controversy about which domain of NTRC causes oligomerization upon phosphorylation. Some reports have placed the oligimerization determinants in the N-terminal receiver domain of NTRC (Fiedler & Weiss, 1995; Mettke et al., 1995). However, a report based on the constitutive mutants described in chapter 5 indicated that the oligomerization determinants are in the central domain (Flashner et al., 1995). The NMR data support this view.

Materials and Methods

Protein Expression and Purification

The protein used in these studies were expressed and purified as described as in the Materials and Methods section of Chapter 3.

Sample Preparation

Two types of samples were prepared. The first method was to flow dialyze concentrated protein solutions against 10 mM phosphate buffer at pH 6.4 and lyophilize.

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 0.1 M HCl or NaOH. Concentrations for both unlabeled and ^{15}N labeled samples varied from 0.25 to 1.5 mM and are indicated in the text. The concentration of the ^{15}N D54N sample was 1 mM. The second method was to flow dialyze concentrated protein solutions against 200 mM phosphate buffer at pH 6.8 and lyophilize. The samples were brought to 0.5 ml in 10% D_2O/90% H_2O or 99.9% D_2O by the addition of the appropriate amounts of H_2O and D_2O. Concentrations for both unlabeled and ^{15}N labelled samples varied from 0.25 to 1.5 mM and are indicated in the text.

Phosphorylation of NTRC

The samples were phosphorylated by the addition of carbamyl phosphate (Sigma) to a final concentration of 100-400 mM followed by the addition of 15-50 mM $MgCl_2$. Alternatively, acetyl phosphate (Sigma) and phosphoramidate were used in place of carbamyl phosphate. Phosphoramidate was synthesized by the method cited (Sheridan et al., 1971).

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.

1D ^1H NMR time courses of phosphorylation and Mg^{2+} binding titrations were performed with a spectral width of 6944 Hz, 8192 total points, the ^1H carrier placed on the H_2O resonance at 4.78 ppm.

1D ^{1}H 1-1 spin echo (Sklenar & Bax, 1987) experiments collected with a spectral width of 6944 Hz, 4096 total points, the ^{1}H carrier placed at 4.78 ppm, and the T_2 relaxation period set to 0.1 ms or 5.1 ms.

^{15}N-^{1}H 2D PEP-Z HSQC (Akke et al., 1994) experiments were collected with spectral widths of 6944 Hz and 2102 Hz in the ^{1}H and ^{15}N dimensions, respectively. The ^{1}H carrier was placed on the H_2O resonance at 4.78 ppm, and the ^{15}N carrier set to 119.1 ppm. A total of 2048 x 128 complex points were collected in the t_1 and t_2 dimensions, respectively. Data were apodized in each dimension with a shifted, skewed sine-bell. A shift of 85° was used in each dimension, with a skew of 1.0 and 0.75 in the t_1 and t_2 dimensions, respectively. Data were zero-filled to yield a 512 x 512 real matrix upon Fourier transformation.

^{15}N-edited 3D NOESY-HMQC (Kay et al., 1989; Marion et al., 1989b) and ^{15}N-edited 3D NOESY PEP-Z HSQC (Akke et al., 1994) experiments were collected with spectral widths of 6944 Hz for the ^{1}H dimensions and 1861 Hz for the ^{15}N dimension. The ^{1}H carrier was placed on the H_2O resonance at 4.78 ppm, and the ^{15}N carrier set to 119.1 ppm. The NOESY mixing time was 80 ms. A total of 96 x 24 x 1024 complex points were collected in the t_1, t_2, and t_3 dimensions, respectively. Data were collected in an interleaved manner due to sample stability restrictions (see the results for details). Data were apodized in each dimension with a shifted, skewed sine-bell. A shift of 85° was used in each dimension, with a skew of 1.0, 0.8, and 0.5 in the t_1, t_2, and t_3 dimensions, respectively. Data were zero-filled to yield a 512 x 64 x 512 real matrix upon Fourier transformation.

Results

Magnesium Binding

Figure 4.1 shows a titration of NTRC receiver domain with $MgCl_2$. The 1D ^{1}H NMR spectrum does not change significantly upon Mg^{2+} binding although a very few small changes in chemical shift can be seen. Unfortunately, NTRC begins to precipitate

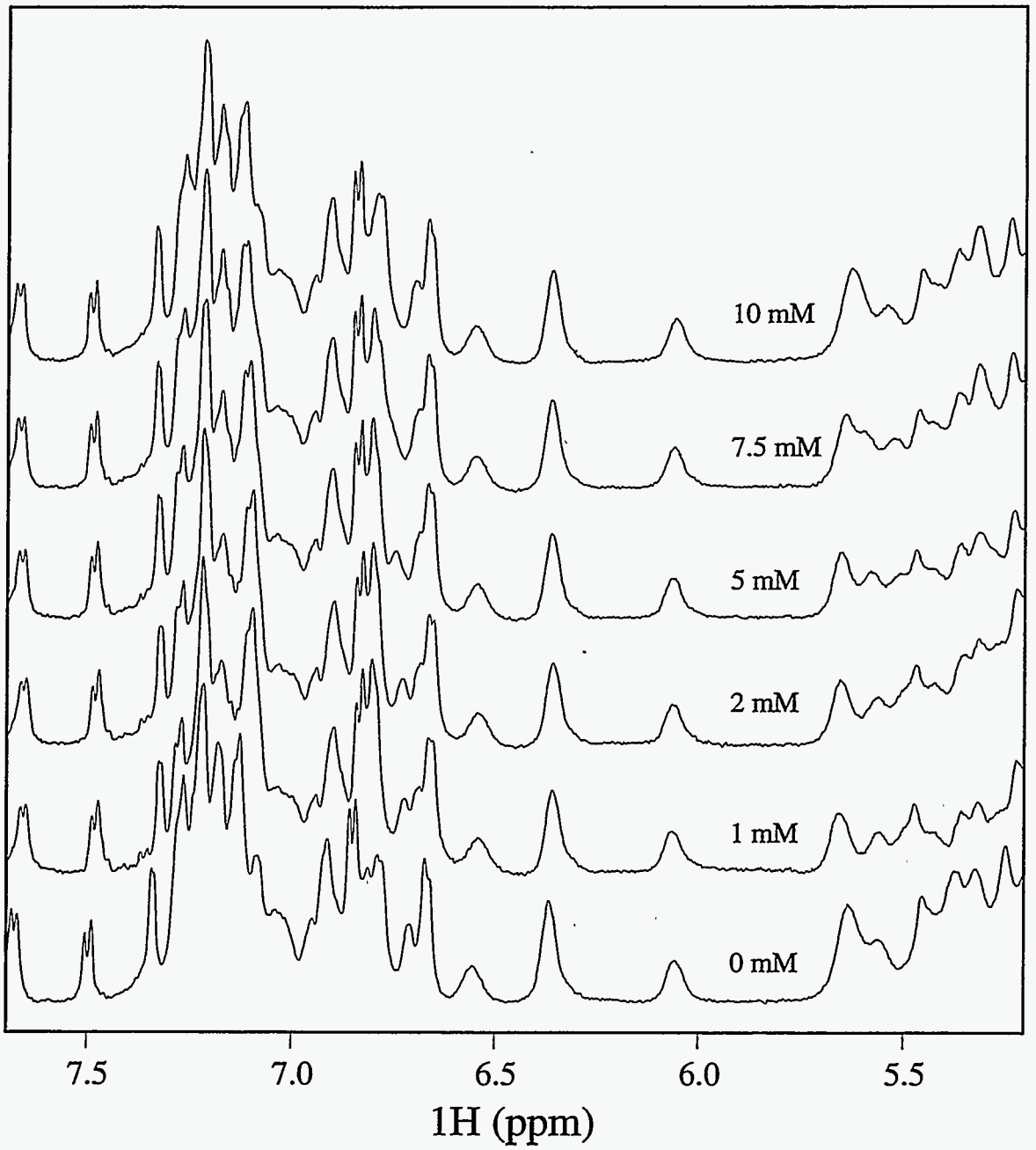


Figure 4.1: Expansion of a ^{1}D ^{1}H NMR spectra showing a titration of 1 mM NTRC receiver domain in 10 mM sodium phosphate pH 6.4 with Mg^{2+}. The concentration of Mg^{2+} is indicated for each spectrum.

upon addition of more than 10 mM $MgCl_2$ which limited the extent of the titration.

Figure 4.2 shows a ^{15}N-^{1}H PEP-Z HSQC spectrum of the Mg^{2+} (10 mM) bound form of NTRC. Tentative assignments are indicated next to each peak. The assignments of the Mg^{2+} bound form were based on assuming a minimum of change from apo-NTRC. This spectrum is very similar to the apo-NTRC spectrum shown in Figure 2.8. Most of the residues make minimal changes upon addition of Mg^{2+}. The notable exceptions to this are the active site residues D11, D12, and D54 which disappear completely at the very lowest concentrations of Mg^{2+}. This is consistent with these residues being involved in the binding of Mg^{2+}. Unfortunately, the precipitation caused by Mg^{2+} made it impossible to complete the titration past 10 mM. It can be rigorously shown that the Mg^{2+} is not saturating at this concentration because the active site residues disappear even when there is less Mg^{2+} (1 mM) than protein (1.5 mM) (data not shown). Interestingly, the residue D10 does not seem to be affected by the addition of Mg^{2+} which indicates that it may not be involved in Mg^{2+} binding.

Phosphorylation of NTRC

There are large rearrangements of the 1D ^{1}H NMR spectrum of NTRC upon phosphorylation. Phosphorylation is achieved by using the small molecule donors acetyl phosphate, carbamyl phosphate, or phosphoramidate. Figure 4.3 shows a portion ^{1}D ^{1}H NMR spectrum of NTRC under phosphorylation conditions (1.0 mM NTRC, 10 mM phosphate buffer pH 6.4, 100 mM carbamyl phosphate, 10 mM $MgCl_2$). The most striking change in the spectrum is the appearance of a new H_α peak at 5.88 ppm. The appearance of this peak was used as a marker of phosphorylation in the determination of optimal conditions. This peak is a particularly good marker of phosphorylation because it is located in a resolved portion of the spectrum and appears, based on intensity, to arise from a single proton.

The phosphorylation reaction, as assayed by the appearance of the peak at 5.88 ppm, is dependent on the presence of Mg^{2+}. Addition of carbamyl phosphate in the

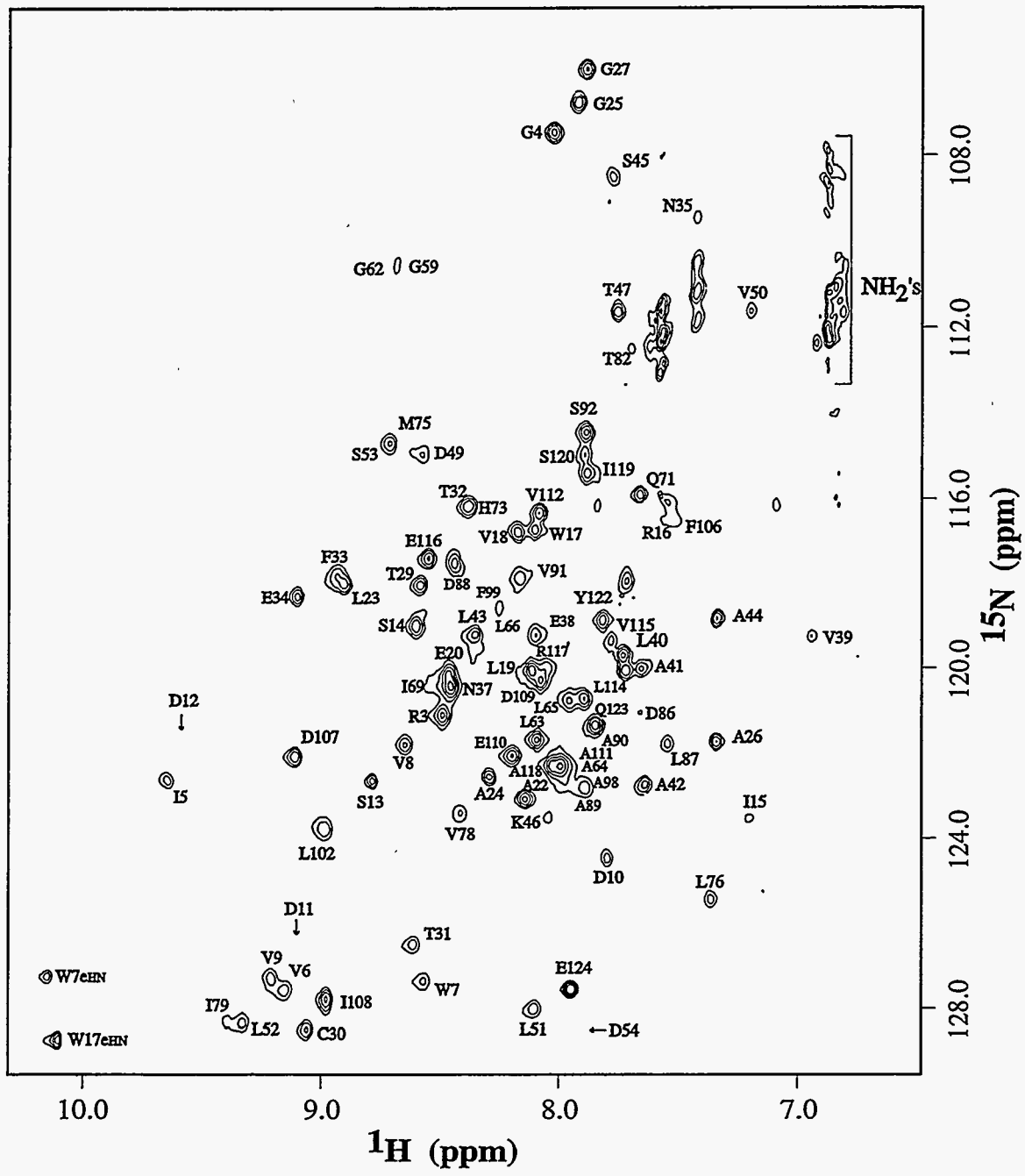


Figure 4.2: ^{1}H-^{15}N HSQC spectrum of the receiver domain of NTRC in the presence of Mg^{2+} (1 mM NTRC, 10 mM sodium phosphate pH 6.4, 10 mM $MgCl_2$). Tentative crosspeak assignments, based on an assumption of minimal difference from apo-NTRC, are labeled. The arrows indicate the position of the active site residues D11, D12, and D54 in apo-NTRC.

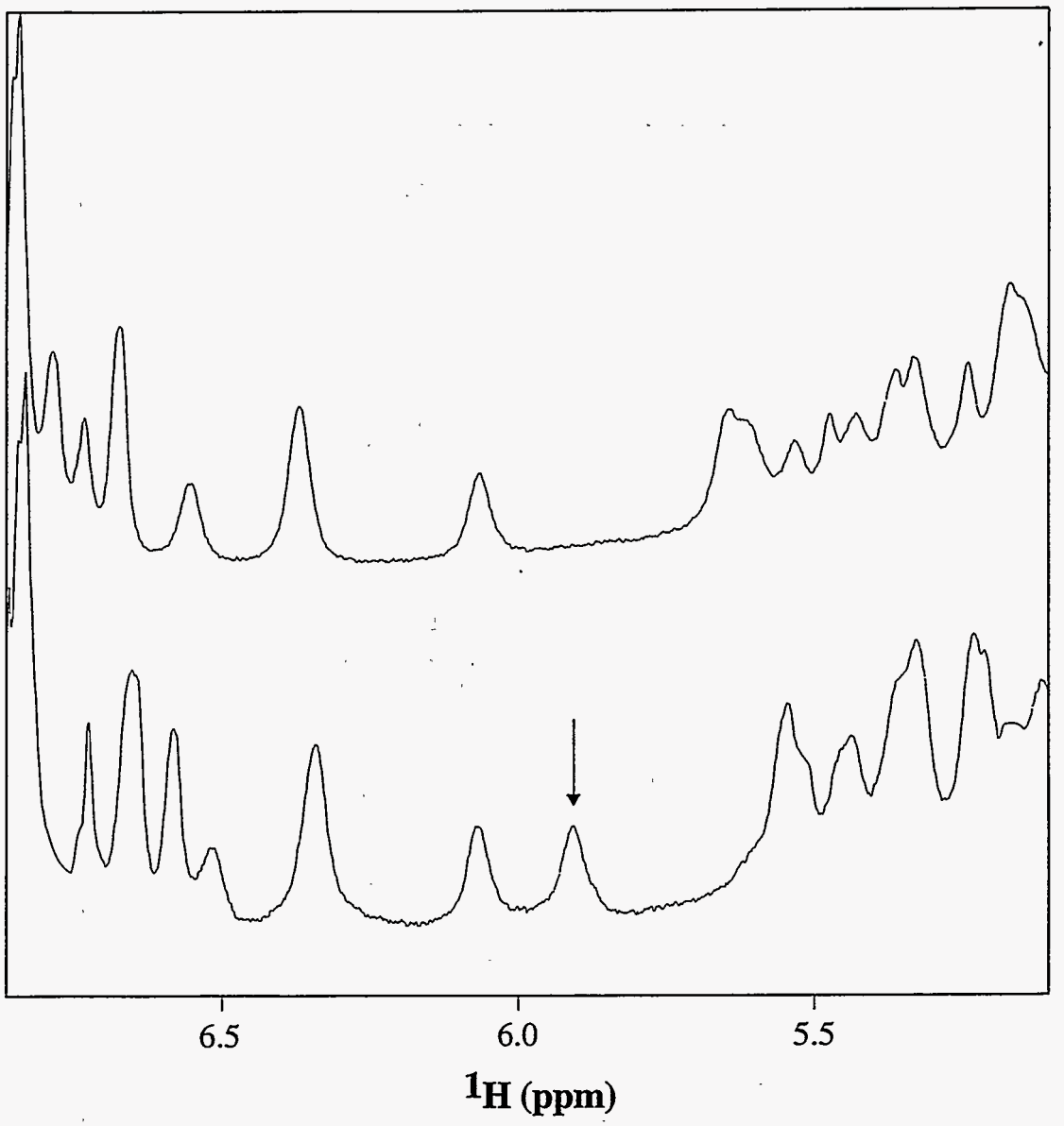


Figure 4.3: Expansion ^{1}D ^{1}H NMR spectra of NTRC in D_2O to follow phosophorylation. The upper spectrum was obtained upon the addition of 100 mM carbamyl phosphate to 1 mM NTRC in 10 mM sodium phosphate buffer, pH 6.4. The lower spectrum shows the effect of the addition of 10 mM Mg^{2+}. The appearance of a single resonance at 5.85 ppm under phosphorylation conditions (indicated by an arow) was used as a marker of phosphorylation.

absence of Mg^{2+} does not result in the rearrangement of the spectrum observed in the phosphorylated form (this is also observed in ^{1}H-^{15}N HSQC spectra of phosphorylated NTRC). This indicates that the changes in the spectra under phosphorylation conditions are not due interactions with the small molecule donor. Furthermore, addition of EDTA to a phosphorylated sample results in the rapid recovery (2-4 min.) of the original spectrum. This implies that although Mg^{2+} is required for the phosphorylation reaction, it is not required for the dephosphorylation of NTRC. Thus, it appears that NTRC's autophosphatase activity is Mg^{2+} independent.

As previously mentioned, 1D ^{1}H NMR spectra were used to probe for optimal phosphorylation conditions. Conditions were evaluated on four criteria. The first criterion was the initial degree of phosphorylation. The second criterion was the length of time the phosphorylation could be maintained. As mentioned above, the phosphorylated state of NTRC has a half-life of about 4 minutes. This is prohibitively short for structural studies. In order to deal with this problem, a steady state of phosphorylation was created by the addition of excess amounts of a phosphodonor. In this scheme, NTRC phosphorylates and dephosphorylates itself many times until the phosphodonor is exhausted. This steady state of phosphorylation could potentially be maintained for many hours depending on the initial concentration of the phosphodonor. The third criterion was the stability of the sample over time. Changes in pH or conformational state in time can lead to severe linebroadening and loss of signal. The fourth criterion was the quality of the spectra, specifically the linewidths of 1D ^{1}H and HSQC spectra.

Acetyl phosphate, carbamyl phosphate, and phosphoramidate were all studied as potential phosphodonors. NTRB was not explored as a phosphodonor because it is not catalytically involved in the phosphorylation. Enormous amounts of phosphorylated NTRB are required to drive the population of NTRC towards the phosphorylated state. The small molecule phosphodonors provide similar phosphodonor capability with much greater solubility. This allows the initial concentration of the phosphodonor to be much

larger and hence allows maintenance of the steady state of phosphorylation for a longer period of time.

Acetyl phosphate only gave about 50% phosphorylation as assayed by the intensity of the peak appearing at 5.88 ppm. Phosphoramidate gave nearly 100% phosphorylation, but this state only persisted for a few hours because phosphoramidate is unstable in aqueous solution. Carbamyl phosphate gave nearly complete phosphorylation (~85-90%) and lasted for up to 15 hours. Based on these properties, carbamyl phosphate was chosen as the phosphodonor of choice for these studies.

Initially, pH was controlled with 10 mM phosphate buffer at pH 6.4. However, it was found that the pH increased up to 1 pH unit over the course of experiments lasting several hours. Most likely, this is due to the conversion of carbamate to ammonia and carbon dioxide. The buildup of ammonia, which has a pKa of 9.2, causes the increase in pH. In order to better control the pH, the concentration of the phosphate buffer was increased to 200 mM and the pH was increased to pH 6.8 which is the pK_a for phosphate buffer. These conditions provide complete pH stability over long time courses (15-18 hr.)

One of the tools used to probe the quality of the spectra was the 1D ^1H 1-1 spin echo experiment (Sklenar & Bax, 1987). This experiment gives a very rough estimate of the T_2 relaxation time of a protein (see chapters 1 and 3). This information can be used as a guide to the suitability of a particular system for study by triple resonance experiments. Triple resonance experiments leave the magnetization in the transverse plane for long periods during multiple constant-time evolution periods. Thus, the length of T_2 is crucial for the success of these experiments. The 1D ^1H 1-1 spin echo is taken once with a T_2 relaxation period of 0.1 ms and once with a period of 5.1 ms and the spectra are compared. If the spectrum taken with the longer T_2 period has less than half the intensity in the amide region of the spectra taken with the shorter T_2 period, then the T_2 for the system is too short (<15 ms) to allow efficient transfers in triple resonance experiments.

Figure 4.4A shows the 1D 1H 1-1 spin echo spectra for unphosphorylated NTRC. In this case, much less than 1/2 of the signal intensity is lost. Figure 4.4B shows the same analysis of NTRC under the best phosphorylation conditions (0.5 mM NTRC, 200 mM phosphate buffer pH 6.8, 300 mM carbamyl phosphate and 27 mM Mg^{2+}). Again, the loss in signal intensity is less than 1/2 indicating the suitability of this system for triple resonance experiments. Also, the loss of intensity in this spectrum is not significantly larger than in the unphosphorylated case indicating that the unphosphorylated and phosphorylated forms of the NTRC have similar T_2's. This, in turn, implies that the two forms have similar molecular weights (see chapter 1).

Figure 4.5 shows a PEP-Z ^{15}N HSQC spectrum of phosphorylated NTRC. This spectrum shows about 85% phosphorylation. This is demonstrated by the peaks corresponding to the W7ε and A42 amides. These residues show peaks from both the phosphorylated and unphosphorylated states (indicated by the arrows). The degree of phosphorylation was determined from the ratio of the phosphorylated intensity to the sum of the phosphorylated and unphosphorylated intensities.

In general, about 50% of the spectrum changes significantly upon phosphorylation indicating some form of conformational change. The signals were tentatively assigned by assuming a minimum of change from the unphosphorylated form of NTRC. The changes in chemical shift with respect to the wild-type form are shown in Figure 4.6. Each bar represents the combined chemical shift differences of ^{1}H and ^{15}N between unphosphorylated and phosphorylated forms of NTRC. This was calculated as follows:

$$\Delta\delta = \sqrt{\left(\Delta^{1}\mathrm{H}\right)^{2} + \left(\Delta^{15}\mathrm{N}\right)^{2}} \qquad (4.1)$$

where $\Delta^{1}H$ and $\Delta^{15}N$ are the chemical shift differences in the proton and nitrogen dimensions, respectively. Negative bars indicate that either the residue was not detectable due to fast exchange or that the assignment for that residue was ambiguous or unavailable

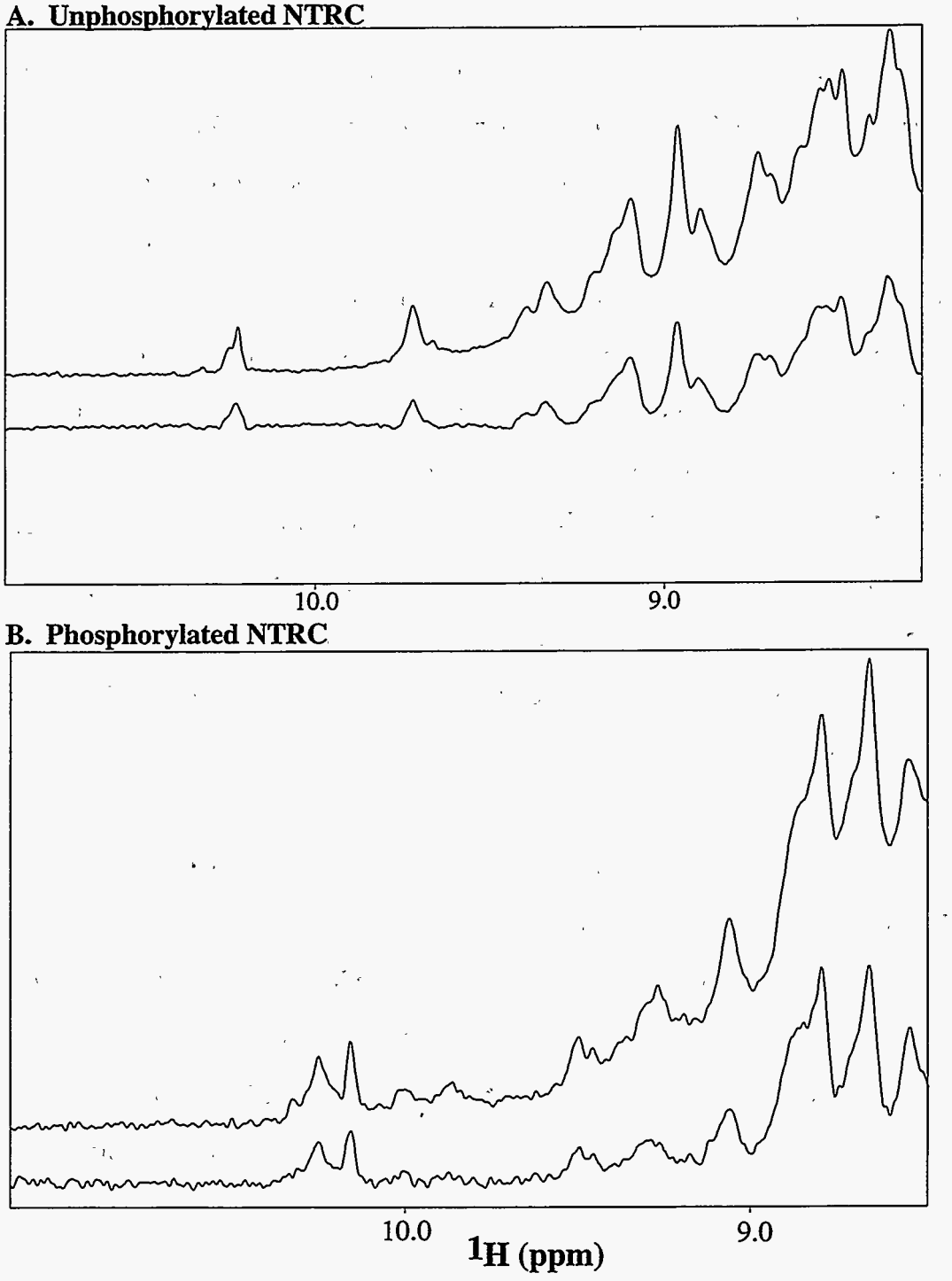


Figure 4.4: 1D ^{1}H 1-1 spin echo spectra of unphosphorylated NTRC (0.5 mM NTRC, 200 mM phosphate buffer pH 6.8) and phosphorylated NTRC (0.5 mM NTRC, 200 mM phosphate buffer pH 6.8, 300 mM carbamyl phosphate, 27 mM Mg^{2+}) with a relaxation period of 0.1 ms (upper spectrum) and 5.1 ms (lower spectrum). Details are given in the text.

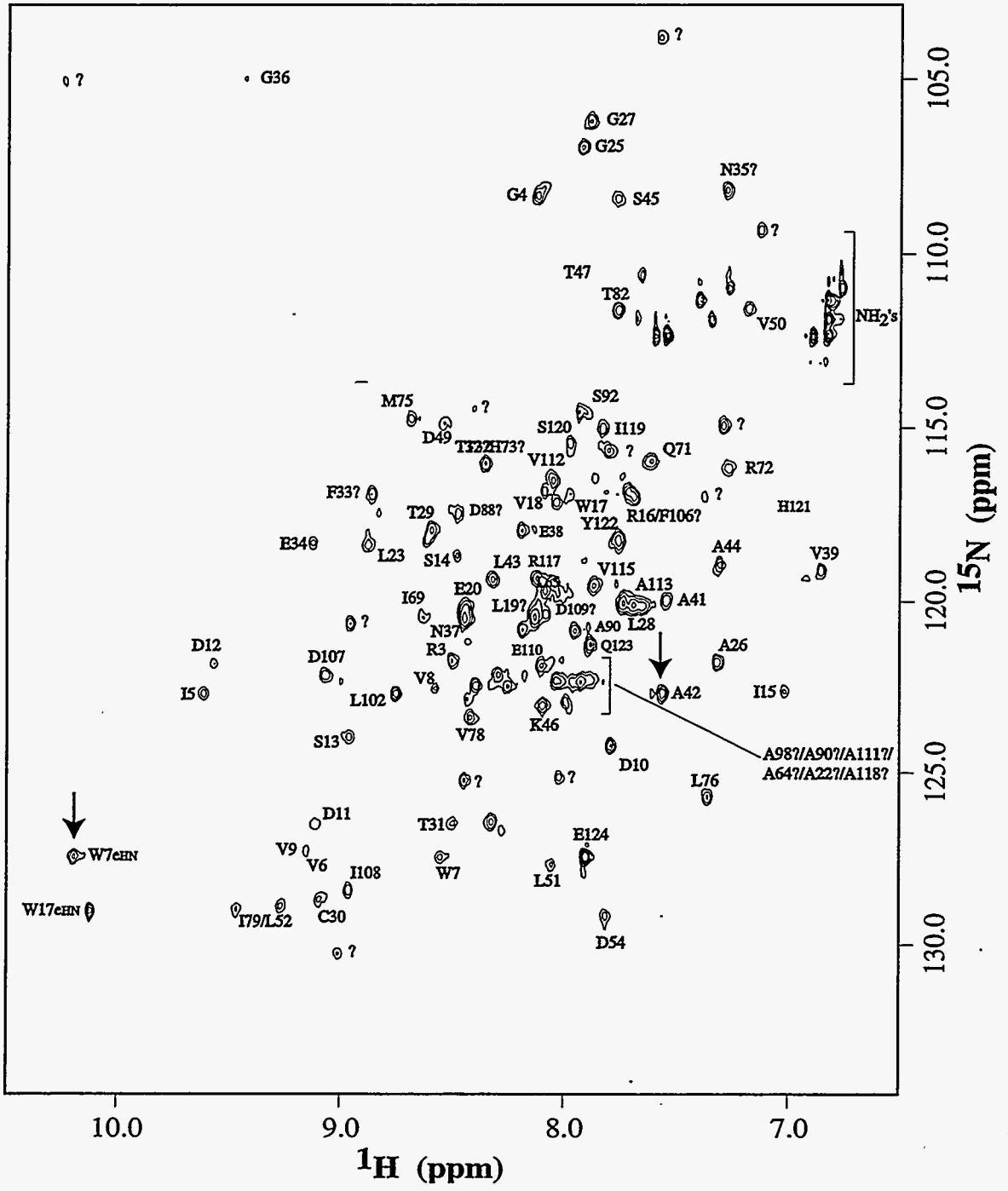


Figure 4.5: ^{1}H-^{15}N HSQC spectrum of the receiver domain of NTRC under phosphorylation conditions (0.5 mM NTRC, 200 mM phosphate buffer pH 6.8, 300 mM carbamyl phosphate and 27 mM Mg^{2+}). Tentative assignments, based on an assumption of minimal difference from unphosphorylated NTRC, are indicated. The arrows indicate peaks used for quantifying the degree of phosphorylation.

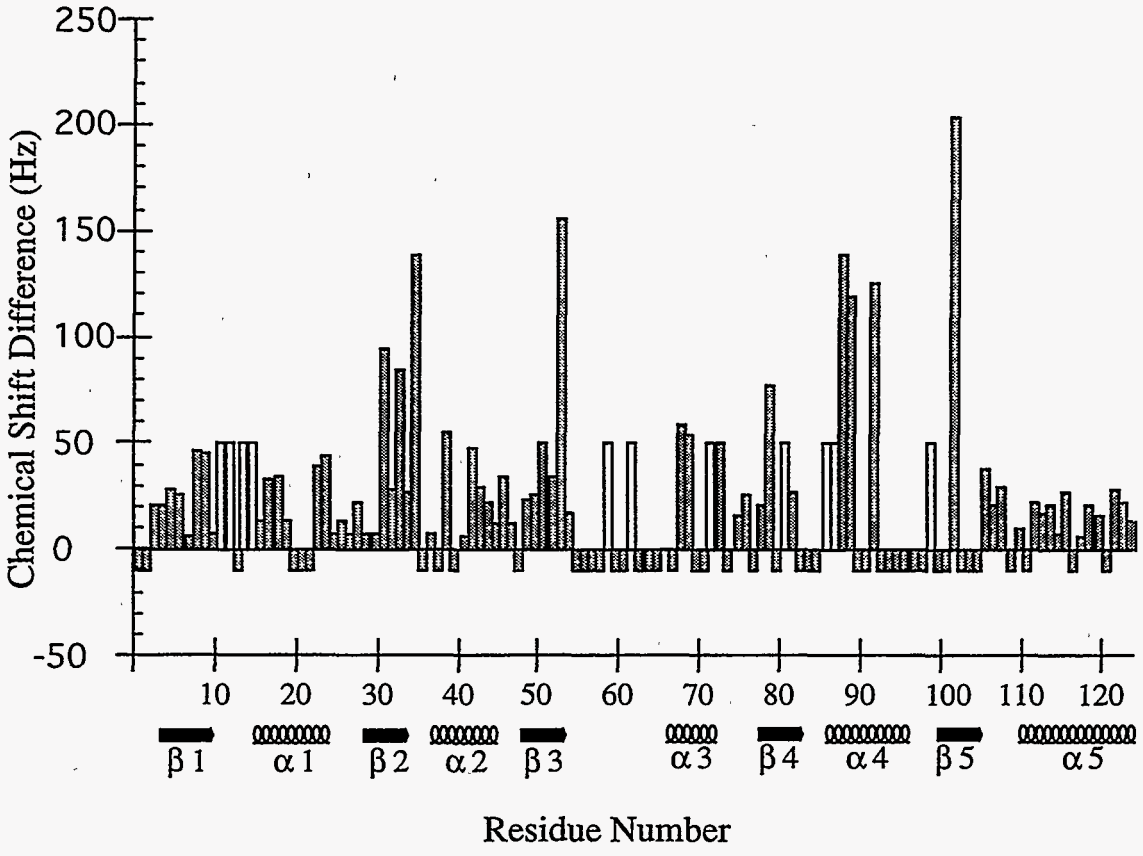


Figure 4.6: Graph of the chemical shift differences between the unphosphorylated and phosphorylated forms of NTRC plotted as a function of residue number. The negative bars represent residues for which either no amide was detectable or the assignments for both forms could not be obtained. Clear bars indicate residues for which large changes are assumed (the peaks have moved so far in the phosphorylated form that they could not be assigned by assuming a minimum of change from the unphosphorylated form).

for the phosphorylated form. Based on these data, phosphorylation causes changes in helix 3, strand 4, helix 4, strand 5 and the top of helix five as well as the area around the site of phosphorylation (the "3445" face of NTRC). This is the same region of the molecule that changes in constitutively active mutants discussed in chapter 5. These changes are mapped onto a structure of the unphosphorylated form of NTRC in Figure 4.7.

The goal of this work is to obtain an atomic resolution structure of the phosphorylated form of NTRC. It requires 30-80 hours to acquire each of the four or five 3 and 4D spectra required for 3-dimensional structure determination by NMR. Thus, conditions need to be maximized to obtain the longest possible equilibrium population of phosphorylated NTRC. The results of a time course on the best set of conditions (0.5 mM NTRC, 200 mM phosphate buffer pH 6.8, 300 mM carbamyl phosphate and 27 mM $MgCl_2$) are shown in Figure 4.8. This figure shows the peak arising from the W7ε imine at varying time points after phosphorylation. At the initial time point, NTRC is about 85% phosphorylated. This falls to about 65% phosphorylated after 15 hours.

One concern was the specificity of the phosphorylation reaction. That is, are the changes seen in the HSQC spectrum due to phosphorylation at D54 and only phosphorylation at D54? In order to test this, a mutant incapable of being phosphorylated, D54N, was prepared. A comparison of HSQC spectra from D54N NTRC alone and under phosphorylation conditions is shown in Figure 4.9. The active site residues are indicated by arrows. There are almost no differences between the spectra. The very slight differences can be attributed to the presence of increased salt concentrations. The active site residues which move significantly in the phosphorylated wildtype spectrum do not move at all in D54N under the same conditions. This indicates that the conformational changes seen in the phosphorylated wildtype spectrum are due to specific phosphorylation on D54. The possibility of phosphorylation on a residue whose amide does not appear in the PEP-Z HSQC spectrum cannot be rigorously excluded. However, if this does occur, it is clearly

Phosphorylated NTRC

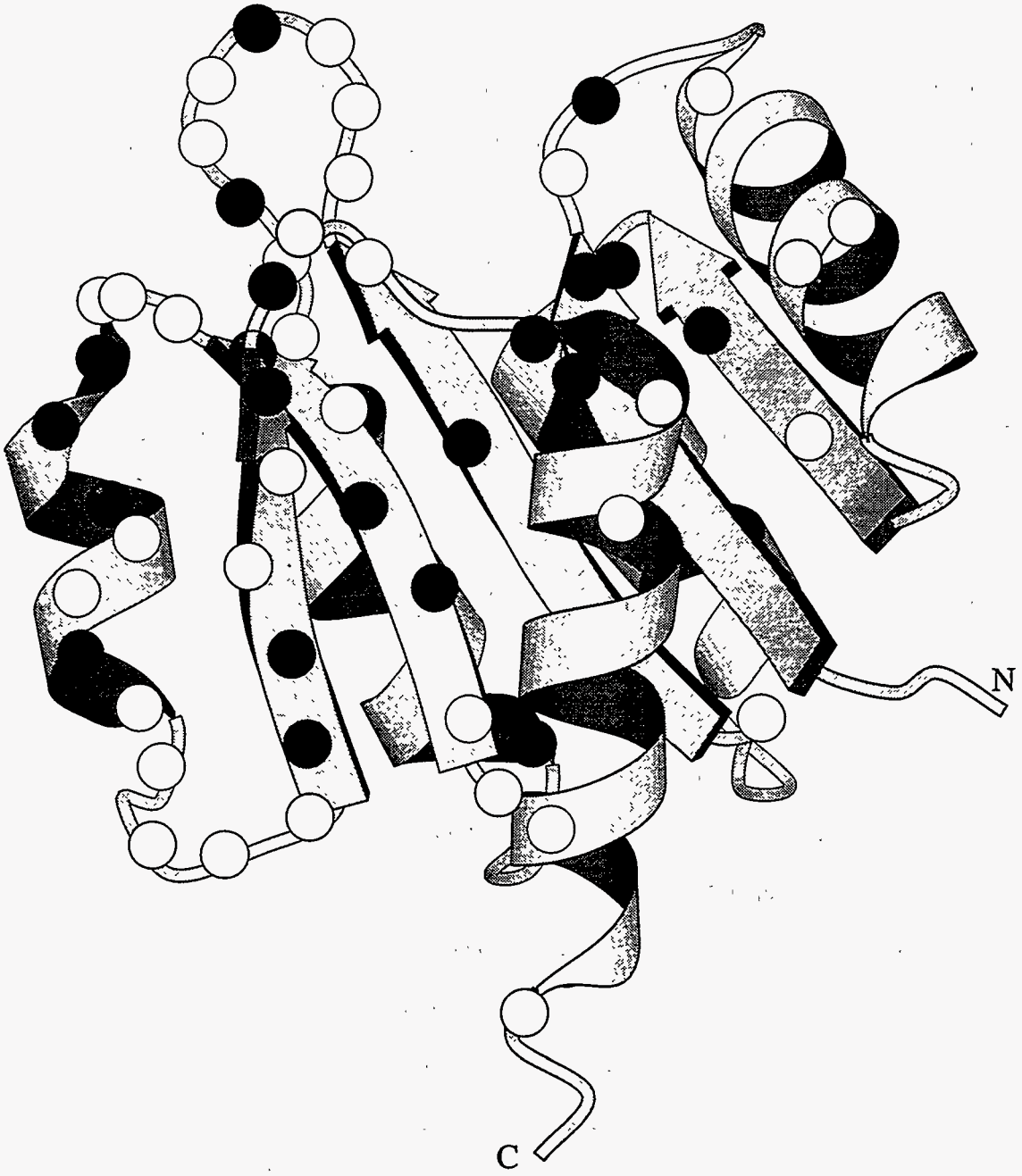


Figure 4.7: The largest chemical shift differences between phosphorylated and unphosphorylated form of NTRC mapped onto the structure of unphosphorylated NTRC. The black spheres represent residues for which the amide resonance have moved at least 50 Hz. This was determined from ^{1}H-^{15}N HSQC spectra. See the text for the details of the calculations. The white spheres represent residues for which either the amide resonances were not detectable or the assignments could not be obtained.

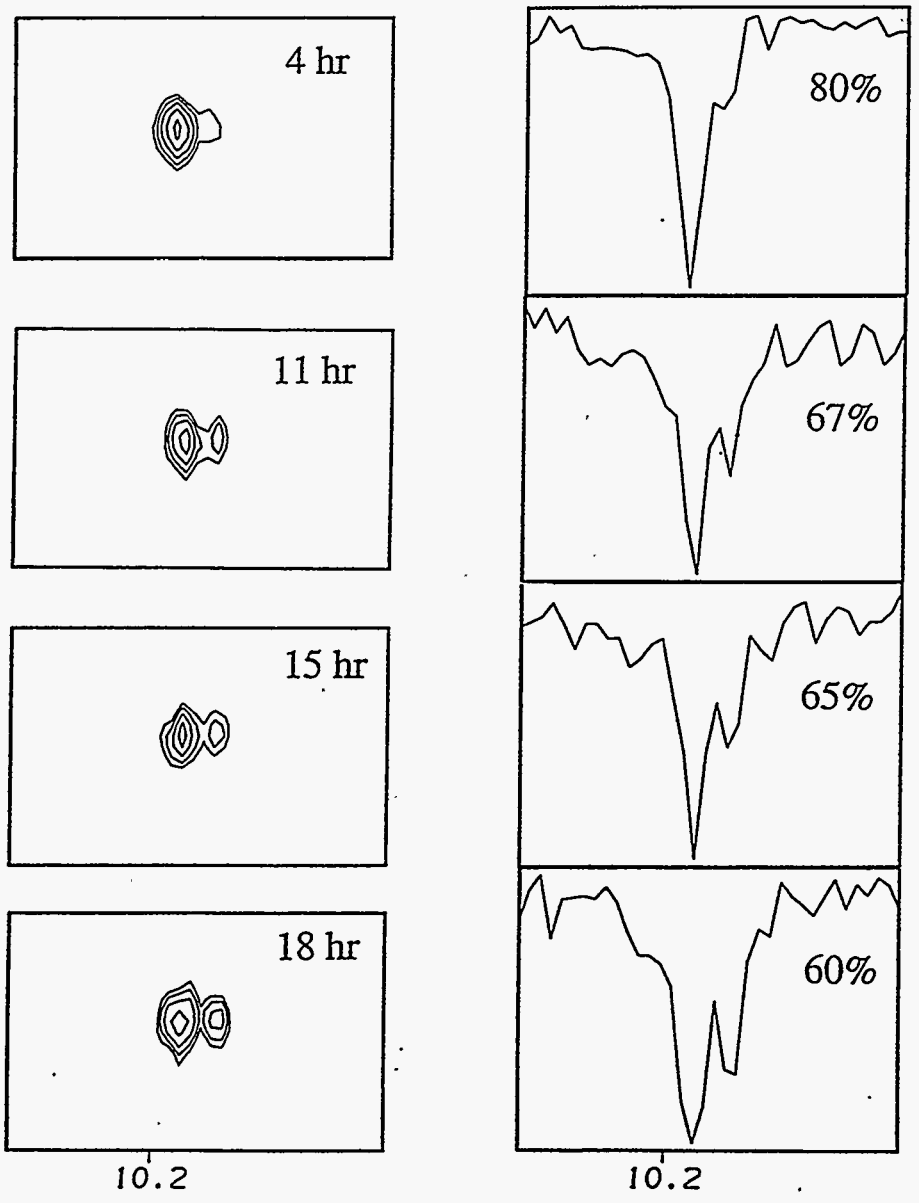


Figure 4.8: A time course of phosphorylation of NTRC. A single peak (W7ε) from a ^{1}H-^{15}N HSQC spectrum of NTRC under phosphorylation conditions (0.5 mM NTRC, 200 mM phosphate buffer pH 6.8, 300 mM carbamyl phosphate and 27 mM Mg^{2+}) is shown. The left column shows an expansion of the ^{1}H-^{15}N HSQC spectrum. The right column shows a 1 dimensional vector taken from the same spectrum. The peak on the left arises from the phosphorylated form of NTRC and the peak on the right arises from the unphosphorylated form of NTRC. The time course is indicated in the left column. The percent of phosphorylation, indicated in the right column, is determined by the ratio of the phosphorylated form to the total amount of signal.

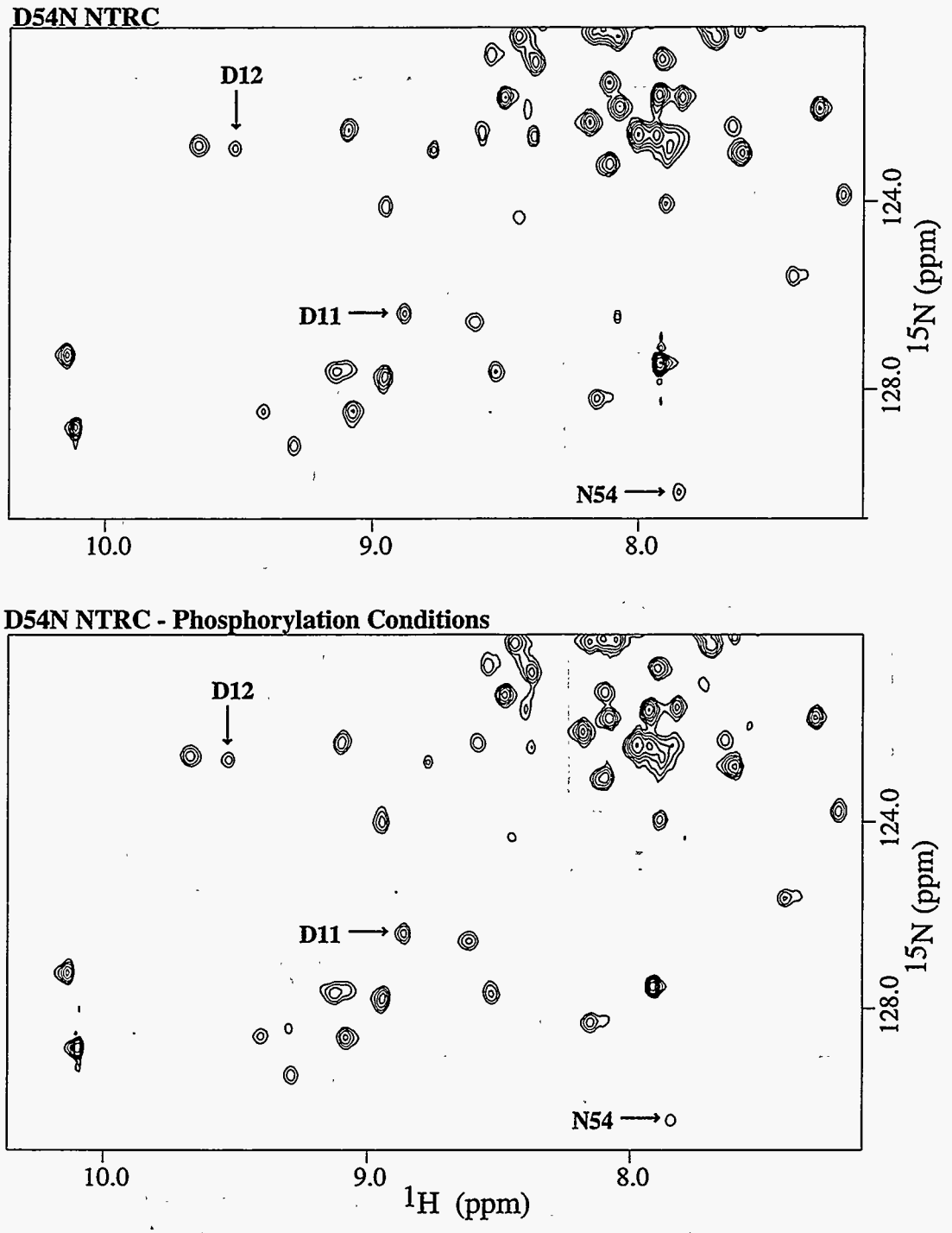


Figure 4.9: Sections of ^{1}H-^{15}N HSQC's of D54N NTRC. The spectrum at the top was taken under non-phosphorylating conditions. The spectrum at the bottom was taken under phosphorylation conditions (the same as Figure 4.5). The lack of significant difference between the spectra indicates that phosphorylation at D54 is specific and responsible for the conformational changes seen in wildtype phosphorylated form of NTRC.

not the cause of the conformational changes detected in the wild-type phosphorylated PEP-Z HSQC spectrum.

Initial Attempts to Assign the Phosphorylated State of NTRC

Initial attempts were made to rigorously assign the backbone of the phosphorylated form of NTRC with an ^{15}N-edited 3D NOESY PEP-Z HSQC and an ^{15}N-edited 3D NOESY-HMQC spectra. Unfortunately, at the time these data were taken, the lifetime of the phosphorylated form was only about 6 hours. This is significantly shorter than the minimum time, about 48 hours, required for acquisition of a 3D NOESY spectrum. In order to obtain a full 3D spectrum, data were taken on eight separate samples and interleaved after acquisition. The dwell time in the ^{15}N dimension of each of the eight experiments was set to eight times the desired dwell time for the reconstructed experiment. The initial dwell delay was set to the desired initial dwell delay times the experiment number (0 to 7). In general, this procedure seemed to recover a complete spectrum. However, an extremely large number of artifacts, including ghost diagonals and severe streaking, were observed. These rendered the spectra uninterpretable. Coworkers have seen the same problems using a similar procedure for 2D NOESY data (Liu, C. and Ho, C., unpublished data).

Discussion

Magnesium Binding to NTRC

Mg^{2+} binding to the N-terminal domain of NTRC appears to cause no change in the conformation of the protein outside of the active site. The disappearance of the peaks from active site residues D11, D12 and D54 could be interpreted in two ways. The most likely explanation is that the Mg^{2+} binding to NTRC is an intermediate exchange process. In this scenario, the peaks disappear due to exchange broadening. This is supported by the fact that the active site residues disappear with even small amounts of Mg^{2+} (~ 1 mM). Another possible explanation is that the active site residues could be in fast exchange and have moved so far by even the lowest concentration of $MgCl_2$ as to be impossible to find.

This seems very unlikely since the active site residues are well separated from the rest of the spectrum. It would require an enormous chemical shift difference between the bound and unbound forms to shift the peaks far enough that they could not be found at low Mg^{2+} concentrations.

It is interesting to note that D10 does not appear to be affected by the presence of Mg^{2+}. This would seem to indicate that it is not involved in Mg^{2+} binding. Unlike the majority of the receiver domain superfamily which have two aspartic acids on the loop between strand 1 and helix 1, NTRC has three (D10, D11 and D12). In the structure of the Mg^{2+} bound form of CheY (Stock et al., 1993), these two aspartic acids (D12 and D13) coordinate the Mg^{2+}. These data indicate that D11 and D12 are the analogous residues which coordinate Mg^{2+} in NTRC.

It was decided not to pursue a Mg^{2+} bound 3D structure of the N-terminal domain of NTRC. The lack of change in the chemical shifts of residues outside the active site upon Mg^{2+} binding makes it unlikely that there are any significant conformational changes to be detected. The changes in the Mg^{2+} bound form appear to be concentrated in the active site which probably cannot be probed due the intermediate exchange phenomena.

Oligomerization of NTRC

The fact that the T_2 relaxation parameters for phosphorylated and unphosphorylated forms of the N-terminal domain of NTRC are very similar (Figure 4.4) indicate that under the conditions of this study, phosphorylated NTRC does not form higher order structures than unphosphorylated NTRC. This follows from the dependence of T_2 on the molecular weight of a molecule. As the size of a molecule increases, the T_2 decreases (see chapter 1). As the phosphorylation conditions involve very large concentrations (0.5 mM) of NTRC, this would seem to support the idea that at least some of the essential dimerization determinants lie outside the N-terminal domain. This supports the work of Flashner et al. (Flashner et al., 1995).

Conformational Change upon Phosphorylation of NTRC

The changes in NTRC upon phosphorylation shown in Figures 4.6 and 4.7 must be interpreted conservatively because the phosphorylated form has not been formally assigned. However, the tentative analysis in Figures 4.6 and 4.7 suggests that a face of NTRC formed by helix 3, strand 4, helix 4, strand 5, and the loop between strand 5 and helix 5 (henceforth this will be referred to as the "3445" face of NTRC) undergo a conformational change upon phosphorylation. This is consistent with the data on the constitutive mutants of NTRC presented in chapter 5.

Progress Towards a Phosphorylated Structure

The conditions presented in Figures 4.5 and 4.8 are quite promising for further characterization of phosphorylated form of NTRC. The largest challenge in such studies is the short half-life of this form. The initial attempts to obtain ^{15}N-edited 3D NOESY spectra failed for two reasons. First, the phsophorylation conditions were suboptimal. The phosphorylated state could only be maintained for about 6 hours and the pH of the sample was drifting during the course of the experiment leading to loss of signal. Second, the interleaved technique used to perform these experiments is prone to artifacts. Recent experience in our laboratory has shown that co-adding rather than interleaving the data leads to superior spectra.

The conditions shown in Figure 4.5 and 4.8 are a great improvement over those used in the ^{15}N-edited 3D NOESY spectra. The pH is stable over the course of the experiment and the phosphorylation is persistent for 18 hours. It would be desirable to increase the initial level of phosphorylation in order to minimize the interference from the minor unphosphorylated form of NTRC. Currently, greater levels of $MgCl_2$ are being investigated as a possible route to this goal. It seems likely that backbone and sidechain assignments could be obtained with the current conditions. However, there is some concern about the ability to derive tertiary structural restraints from a ^{13}C edited NOESY experiment. If the dynamic range of NOE intensities in such a spectrum is an order of

magnitude, then the most intense NOEs from the minor form present at 15% could appear as strong as the weak NOEs from the major form.

Unfortunately, the Bruker AMX-600 on which the ^{15}N-edited 3D NOESY spectra were taken is incapable of acquiring a 3D spectra quickly enough to make co-adding a viable strategy. However, on a spectrometer that can use pulsed-field gradients for the selection of coherence pathways (Kay, 1995) this should be feasible. Indeed, some triple resonance experiments may require as little as 24 hours with this sort of spectrometer obviating the need for multiple samples.

Summary

It has been shown that there is not a significant change in the structure of NTRC outside of the active site upon Mg^{2+} binding. Conditions for structural studies of the phosphorylated form of NTRC have been obtained. Preliminary analysis suggests that a significant portion of the molecule, encompassing helix 3 the loop between strand 5 and helix 5 (the "3445" face of NTRC), undergoes a conformational change upon phosphorylation.

Chapter 5 The Activated Form of NTRC via Mutational Analysis

Introduction

Overview

Chemical shift differences between the amides of constitutively active mutants and wildtype NTRC have been determined by ^{1}H-^{15}N HSQC experiments. These changes are mapped onto the wild type NTRC structure to suggest the regions of the protein involved in activation. Finally, they are compared to changes that occur in the phosphorylated form of NTRC.

Constitutive Mutations of NTRC

Constitutively active mutants of NTRC, those which stimulate transcription without phosphorylation, provide a stable form of the molecule in the activated state. This offers a route to the analysis of the active NTRC without having to deal with the technical difficulties associated with the phosphorylated form (see chapter 4). Kustu and coworkers identified a number of mutants of NTRC that are constitutive for activation (Flashner et al., 1995). That is, the mutated proteins stimulate transcription without being phosphorylated. These mutants were isolated by selecting for suppression of null alleles of NTRC's cognate histidine kinase NTRB. Normally, strains lacking NTRB activity will not grow with arginine as the sole nitrogen source. This strategy identified a number of constitutive mutants in both the N-terminal receiver domain and the central activation domain. The mutants in the N-terminal receiver domain, D86N and A89T, map to the top of helix 4. More recently, another constitutive mutant, V115I was found by a similar selection strategy. This mutant maps to the middle of helix 5. All of these mutants have been shown to activate transcription *in vitro*. One other constitutive mutant, D54E, was made by site-directed mutagenesis. It was designed by attempting to mimic the effect of phosphorylation on D54. Changing the aspartic acid to a glutamic acid moves the carboxyl charge further out into the active site pocket. If the phosphorylation causes its effect through electrostatic interactions, then this mutation will cause a similar, if smaller, effect.

Indeed, D54E had the expected phenotype and is capable of activating transcription *in vitro*, but with low activity. In general, the single constitutive mutants are quite low in activity compared to wild-type phosphorylated NTRC. In *in vitro* transcription assays, the order of activity of the single mutants from lowest to highest is V115I < D54E < D86N < A89T (Figure 5.1A). The activity of V115I is negligible which is puzzling in light of its ability to survive the selection described above.

In order to characterize the mutants further, they were made in combination (Stedman, K. and Kustu, S., personal communication). The activities of the mutants are significantly more than additive. The double mutant D86N/A89T is several times more active than the best single mutant, A89T. Surprisingly, the triple mutant, D86N/A89T/V115I, is somewhat better than D86N/A89T (Figure 5.1B) whereas the double mutants containing V115I are not much better than the single mutants (data not shown). The mutants have also been compared to the activity of wild-type phosphorylated NTRC and the strongest mutant, D86N/A89T/V115I, has about 50% of the activity of wildtype (Figure 5.1C).

The above results suggested that it would be interesting to investigate the structural basis of the activation of the constitutive mutants. This is particularly attractive since constitutive mutants in the active site (D54E) and outside of the active site (D86N, D86N/A89T, D86N/A89T/V115I) could be compared. Of particular interest is the structural basis of the additivity of activities of the mutants. Also, the structural changes in the constitutive mutants can be compared to the changes in the phosphorylated wildtype form of NTRC. Finally, since the constitutive mutants do not require phosphorylation for activity, they were also useful for the determination of the role of Mg^{2+} in the activation of NTRC.

Mutations in CheY

Studies of the mutations in CheY that can be suppressed by compensatory mutations in the flagellar switch proteins have attempted to define an interaction surface

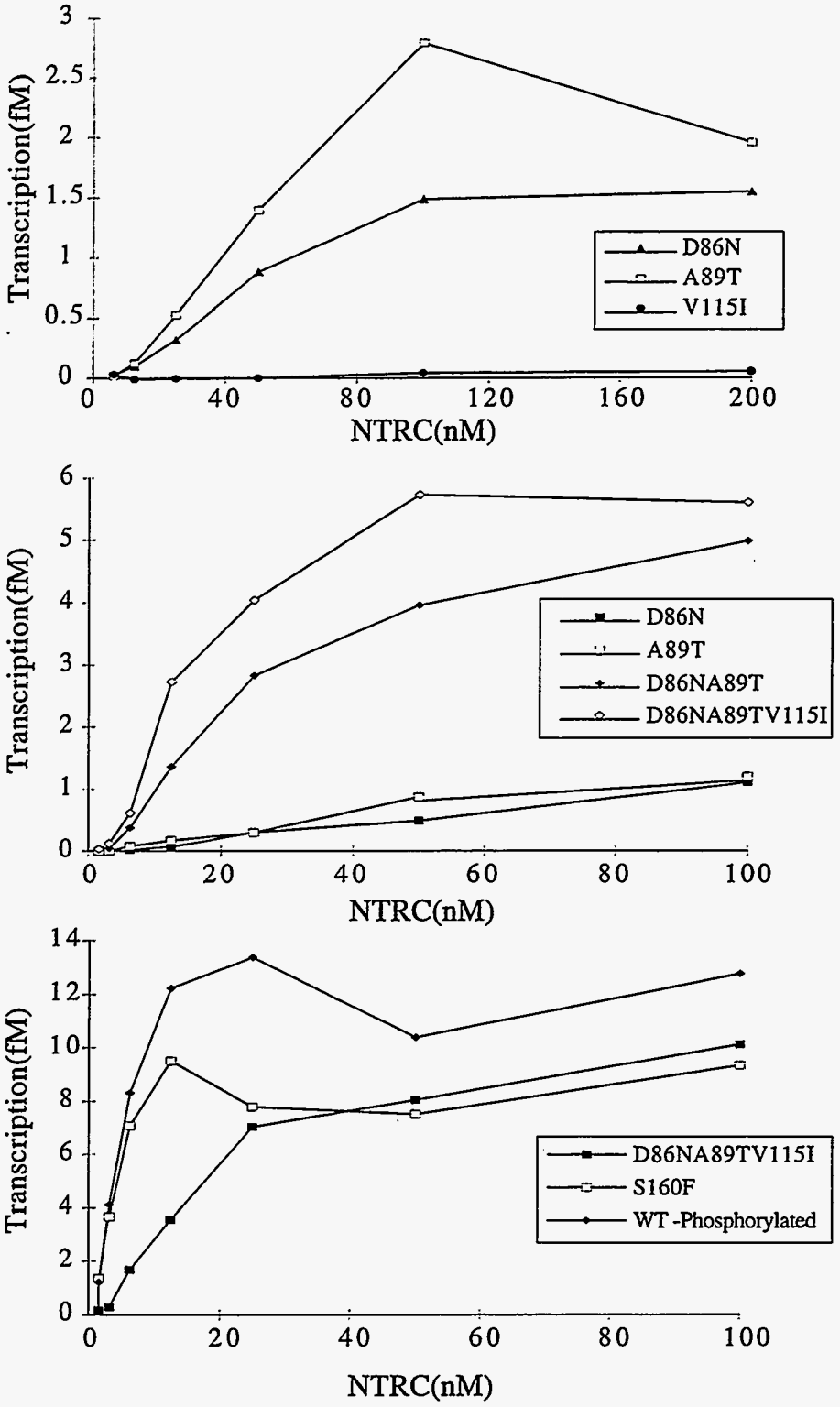


Figure 5.1: *In vitro* transcription assays of NTRC mutants and wildtype phosphorylated NTRC. All assays were performed with full length NTRC. The various mutants are indicated in the figure legends. This figure was kindly provided by Ken Stedman in Sydney Kustu's laboratory.

(Roman et al., 1992; Sockett et al., 1992). The suppressible mutations in CheY are spatially clustered in helix 4, strand 5, the loop between strand 5 and the top of helix 5. This suggests that this region of the protein may be an interaction surface with downstream target flagellar proteins. It is interesting to note that the area of CheY that appear to change upon phosphorylation (Lowry et al., 1994), encompassing helix 3 to helix 5, is larger than the area that contains suppresser mutations.

A structure for another interesting mutation in CheY, T87I, has been obtained to a 2.1 Å resolution (Ganguli et al., 1995). In this mutant, phosphorylation fails to lead to activation. Overall, this mutation does not perturb the structure except for small local perturbations near T87I itself. T87 is an active site residue in CheY that corresponds to T82 in NTRC. Thus, this study suggests that T87 in CheY (and hence T82 in NTRC) may be important for the propagation of the conformational changes required for activation upon phosphorylation.

Material and Methods

Expression and Purification

For the N-terminal domain of NTRC (residues 1-124) the expression vector pJES592 (Klose et al., 1994), which includes a T7 promoter, was used. The expression vectors for the mutants, the details of which can be found in the reference cited (Stedman, 1996), contained the sequence for the mutant NTRC and a T7 promoter. These expression vectors were transformed into *E. coli* BL21 (DE3) cell carrying the pLysS plasmid (Studier et al., 1990).

The wildtype and mutant versions of NTRC were expressed and purified as described in Material and Methods of chapter 3.

Sample Preparation

Samples were prepared as described in Material and Methods of chapter 4 with special attention to matching the conditions in all the samples. This was accomplished by performing flow dialysis against the same buffer at the same time.

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 (^{1}H 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-^{1}H 2D PEP-Z HSQC (Akke et al., 1994) experiments were collected with spectral widths of 6944 Hz and 2102 Hz in the ^{1}H and ^{15}N dimensions, respectively. The ^{1}H carrier was placed on the H_2O resonance at 4.78 ppm, and the ^{15}N carrier set to 119.1 ppm. A total of 2048 x 128 complex points were collected in the t_1 and t_2 dimensions, respectively. Data were apodized in each dimension with a shifted, skewed sine-bell. A shift of 85° was used in each dimension, with a skew of 1.0 and 0.75 in the t_1 and t_2 dimensions, respectively. Data were zero-filled to yield a 512 x 512 real matrix upon Fourier transformation.

^{15}N-edited 3D NOESY-HMQC (Kay et al., 1989; Marion et al., 1989b) experiments were collected with spectral widths of 6944 Hz for the ^{1}H dimensions and 2102 Hz for the ^{15}N dimension. The ^{1}H carrier was placed on the H_2O resonance at 4.78 ppm, and the ^{15}N carrier set to 119.1 ppm. The NOESY mixing time was 100 ms. A total of 128 x 32 x 1024 complex points were collected in the t_1, t_2, and t_3 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 t_1, t_2, and t_3 dimensions, respectively. Data were zero-filled to yield a 512 x 64 x 512 real matrix upon Fourier transformation.

Results

Magnesium and the Activated State of NTRC

Figure 5.2 shows the PEP-Z HSQC spectrum of D54E in the presence and absence of $MgCl_2$. In general, no changes can be observed between the spectra. Of special interest is the fact that the active site residues, indicated with arrows, do not move at all in the presence of Mg^{2+}. Note that active site residues, D11, D12, and D54 do not change chemical shift in the presence of Mg^{2+} implying that the mutant does not bind Mg^{2+}. This is in direct contrast with the situation in wild-type NTRC where the active site residues completely disappear in the presence of Mg^{2+}. This is most likely due to an intermediate exchange line broadening effect (see chapter 4). Thus, it appears that Mg^{2+} binding is required for phosphorylation, but not for activation of the molecule. This was fortunate because it allowed the studies of the constitutive mutants to be carried out in the absence of Mg^{2+} which causes precipitation problems.

Comparison of HSQC Spectra of Mutant and Wildtype NTRC

HSQC spectra of all mutants have been acquired and compared to the HSQC of wildtype NTRC. Figure 5.3 and Figure 5.4 shows a comparison of a wildtype NTRC PEP-Z HSQC spectrum with a PEP-Z HSQC spectrum of D86N/A89T/V115I NTRC as an example. The mutant peaks were tentatively assigned by assuming a minimum change from the wildtype spectrum. These initial assignments were subsequently confirmed with the ^{15}N edited 3D NOESY-HMQC experiment which also provided some new assignments. Overall, many of the peaks (about 60%) don't move significantly (<20 Hz) indicating that these residues have not undergone structural rearrangement. The extensive shifts in the other residues suggest some form of structural rearrangement has occurred. The other mutants, D54E, V115I, D86N and D86N/A89T, were assigned on the basis of the assignments of the triple mutant and on the assumption of minimum changes form the wildtype spectrum.

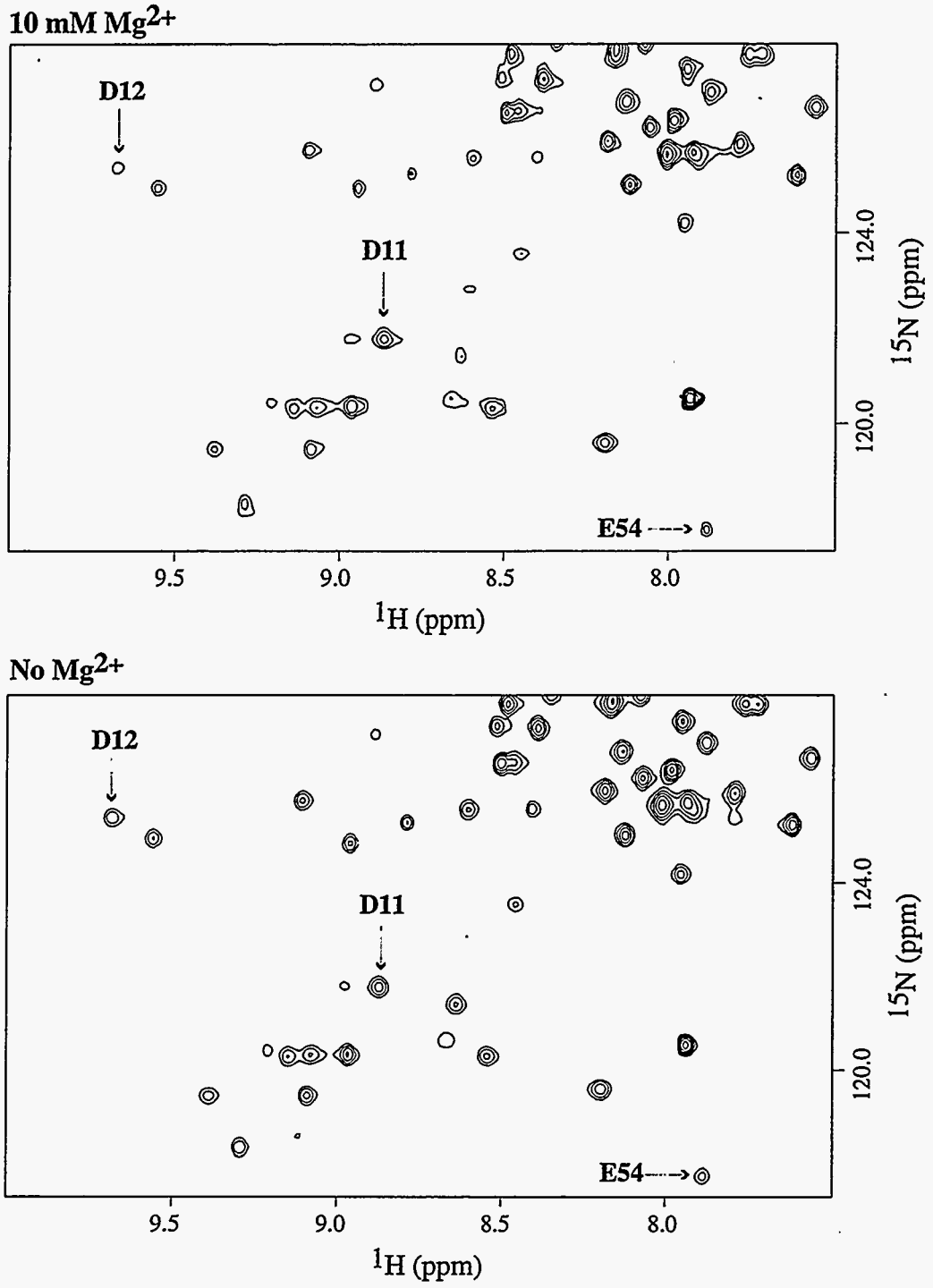


Figure 5.2: Sections of ^{1}H-^{15}N HSQC's of D54E NTRC in the presence and absence of Mg^{2+}. The positions of the active site residues D11, D12, and E54(D54) are indicated by arrows. The lack of change in these residues argues that D54E NTRC does not bind Mg^{2+} at the active site.

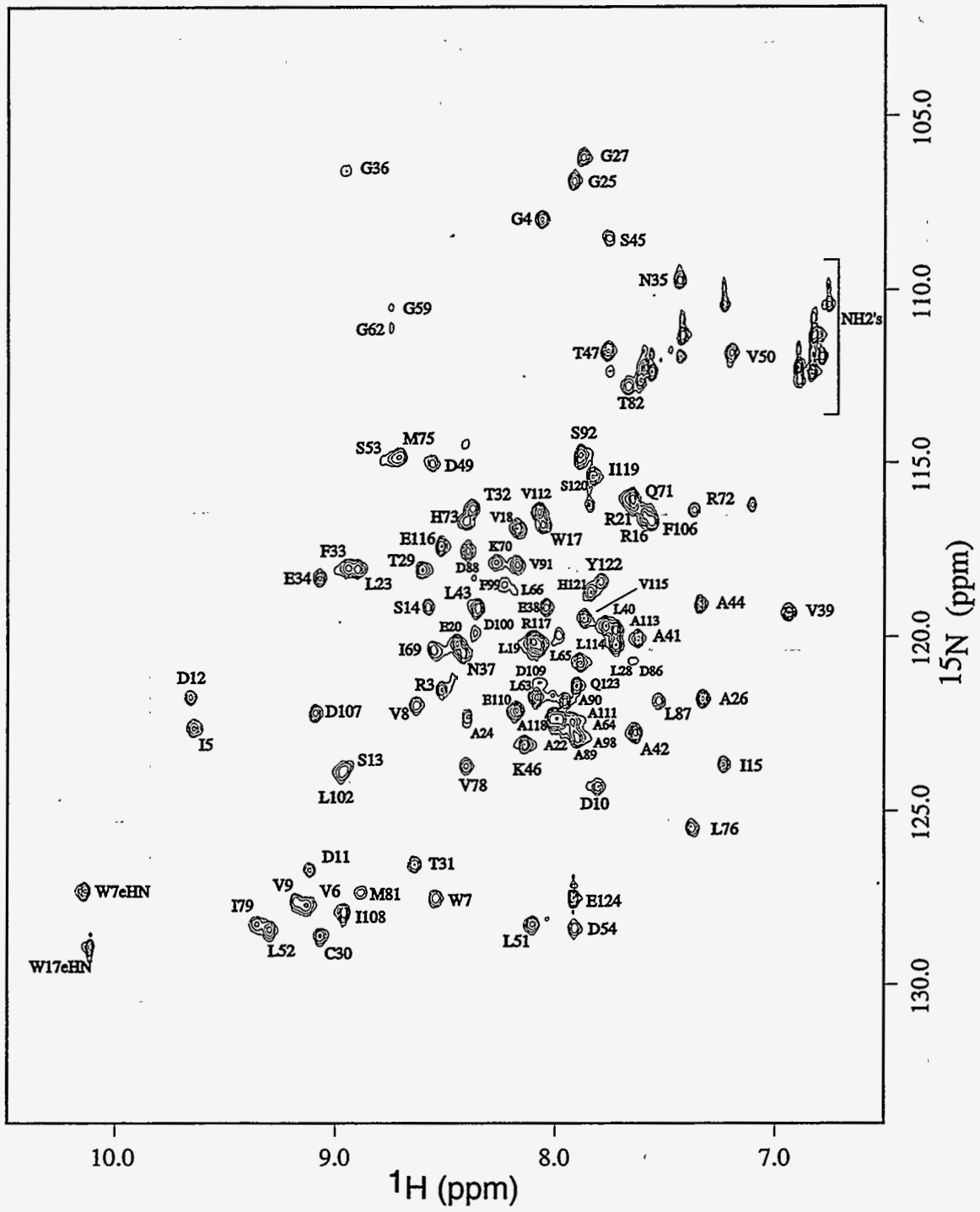


Figure 5.3: ^{1}H-^{15}N HSQC of wild type NTRC (200 mM phosphate buffer pH 6.8). Assignments are indicated by the one letter amino acid codes and residue number. Note that there are very small rearrangements in this spectrum compared to that shown in Figure 2.8 due to differences in pH and buffer concentration.

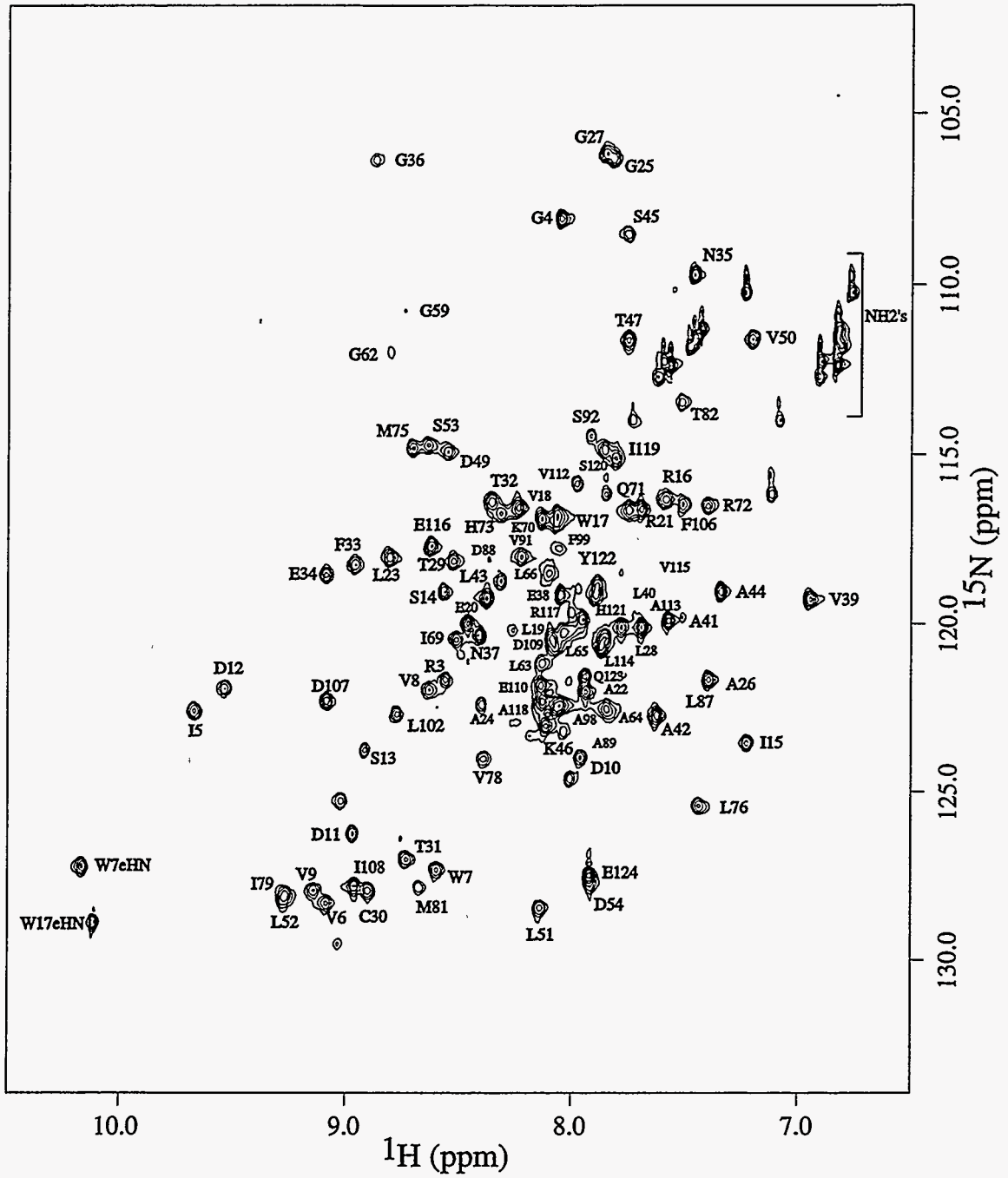


Figure 5.4: ^{1}H-^{15}N HSQC of D86N/A89T/V115I NTRC (200 mM phosphate buffer pH 6.8). Assignments are indicated by the one letter amino acid codes and residue number.

The chemical shift differences between the amides of the mutant and wildtype NTRC were calculated as follows:

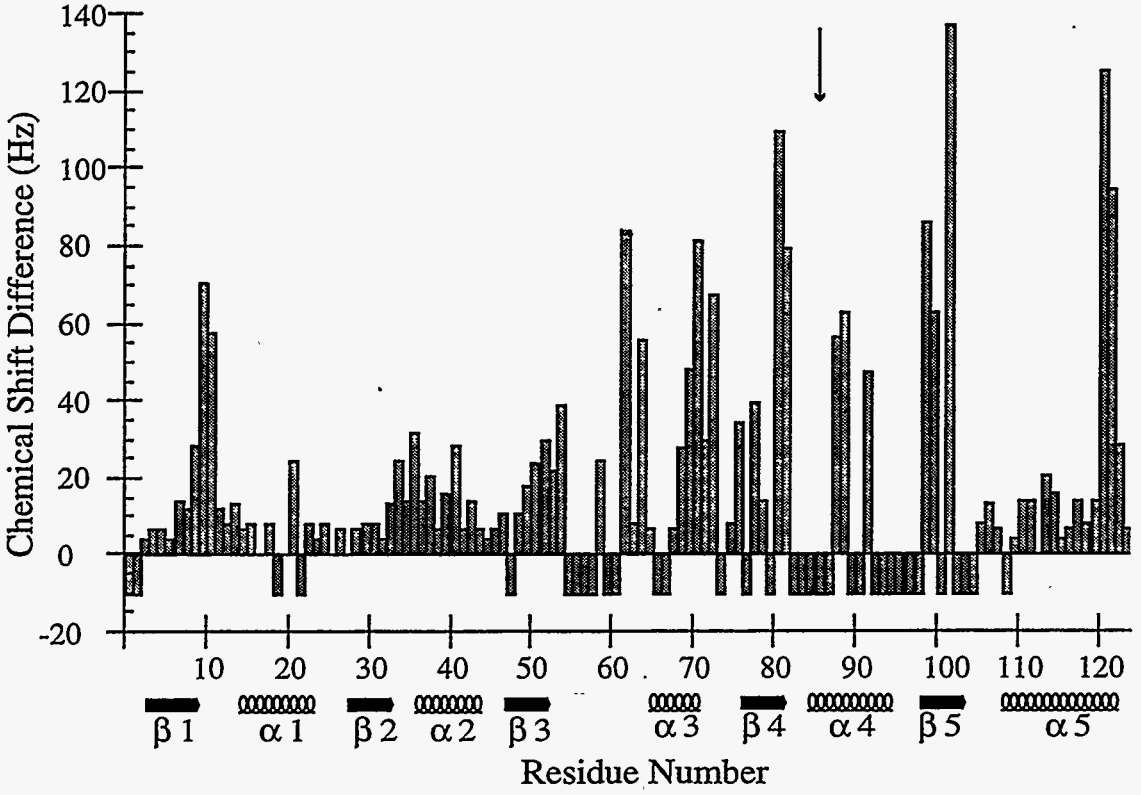
$$\Delta\delta = \sqrt{\left(\Delta^{1}\mathrm{H}\right)^{2} + \left(\Delta^{15}\mathrm{N}\right)^{2}} \qquad (5.1)$$

where $\Delta^{1}\mathrm{H}$ and $\Delta^{15}\mathrm{N}$ are the chemical shift differences in the proton and nitrogen dimensions, respectively. Residues for which either or both the wild type and mutant assignments were unavailable are indicated by a negative bar. The positions of the mutated residues are indicated with arrows.

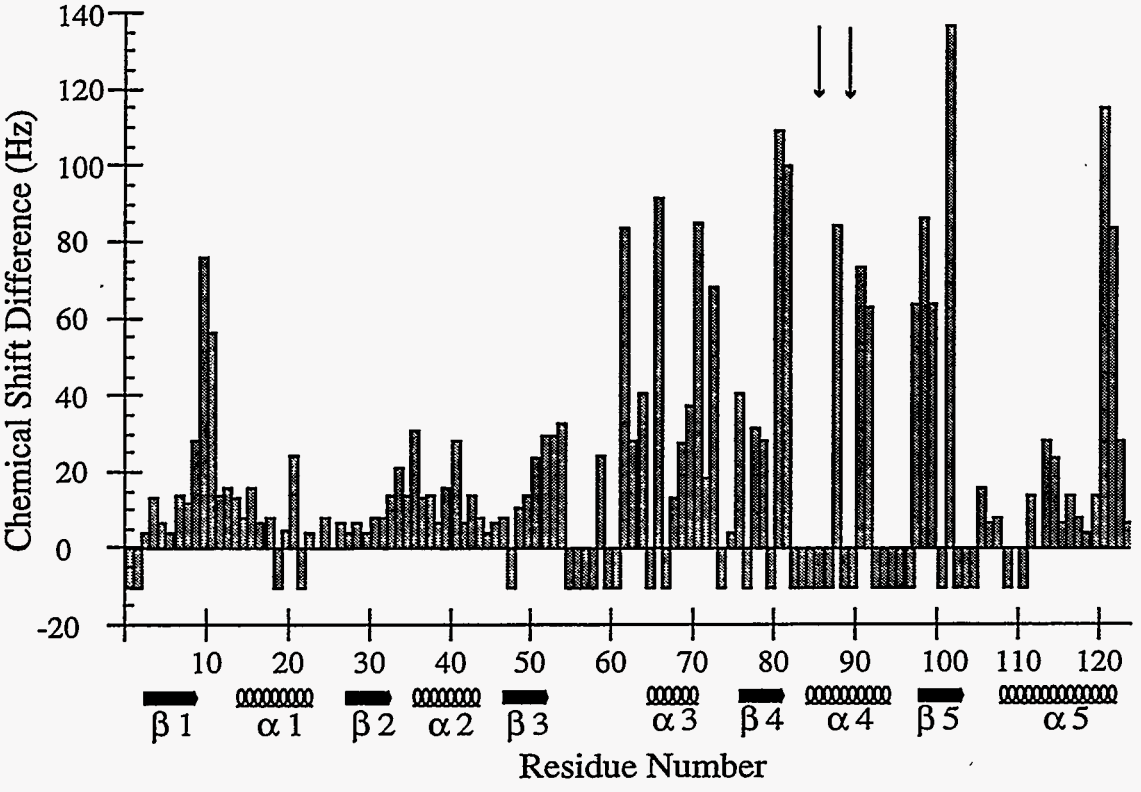
The residues showing significant chemical shift change in D86N NTRC, Figure 5.5A, are found in the loop between strand 1 and helix 1 and the loop between strand 5 and helix 5. These regions are part of the active site or very close to it. In addition, there are large changes in helix 4 (which are expected as this is the site of mutation), the end of helix 3, strand 4, and strand 5. The residues showing the largest chemical shift changes in the double mutant D86N/A89T NTRC, Figure 5.5B, lie in essentially the same regions of the molecule as those in D86N NTRC alone although some of the changes are larger. The residues showing significant changes in the triple mutant D86N/A89T/115I, Figure 5.5C, are similar to those of D86N/A89T NTRC, although again larger. In addition changes are seen in helix 1, strand 2, and helix 5. These residues are in proximity to the site of the additional mutation, V115I. To prove the effect of the mutation at V115I on this area of the protein, analysis of the V115I (Figure 5.5D) single mutant was undertaken even though the activity is very low. In fact, changes were found as expected in helix 1, strand 2, and helix 5. Interestingly, also the active site residues D11 and D12 move significantly.

In addition to these non-active site mutants, the active site constitutive mutant D54E was investigated. The residues showing significant chemical shift changes in D54E NTRC, Figure 5.5E, are clustered in the active site and in the top of strand 4, helix 4 and strand 5.

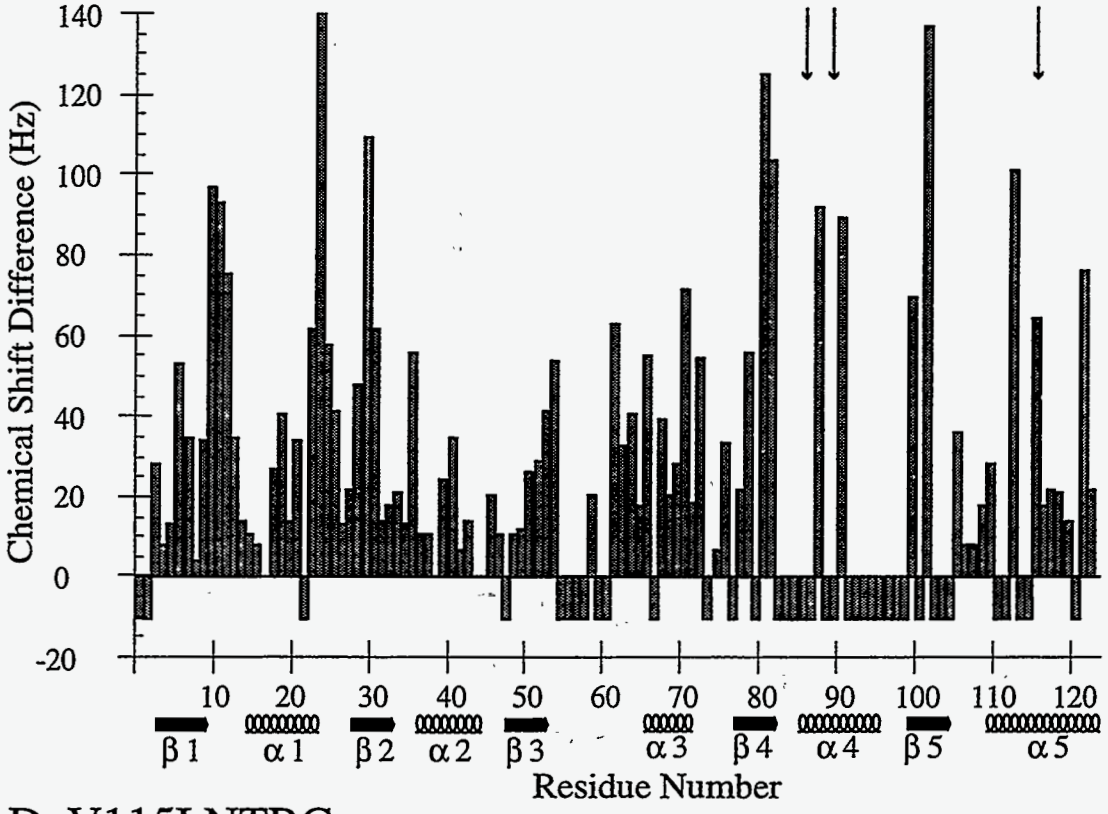
A D86N NTRC



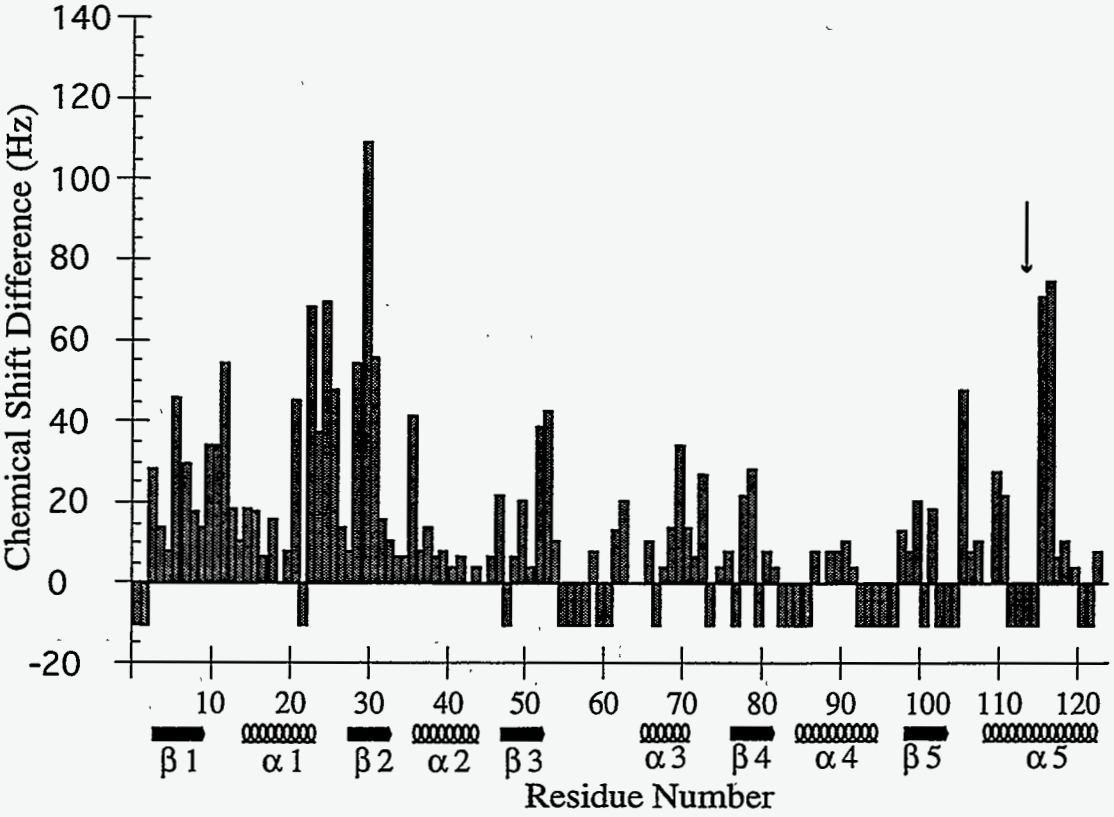
B D86NA89T NTRC



C D86N/A89T/V155I NTRC



D V115I NTRC



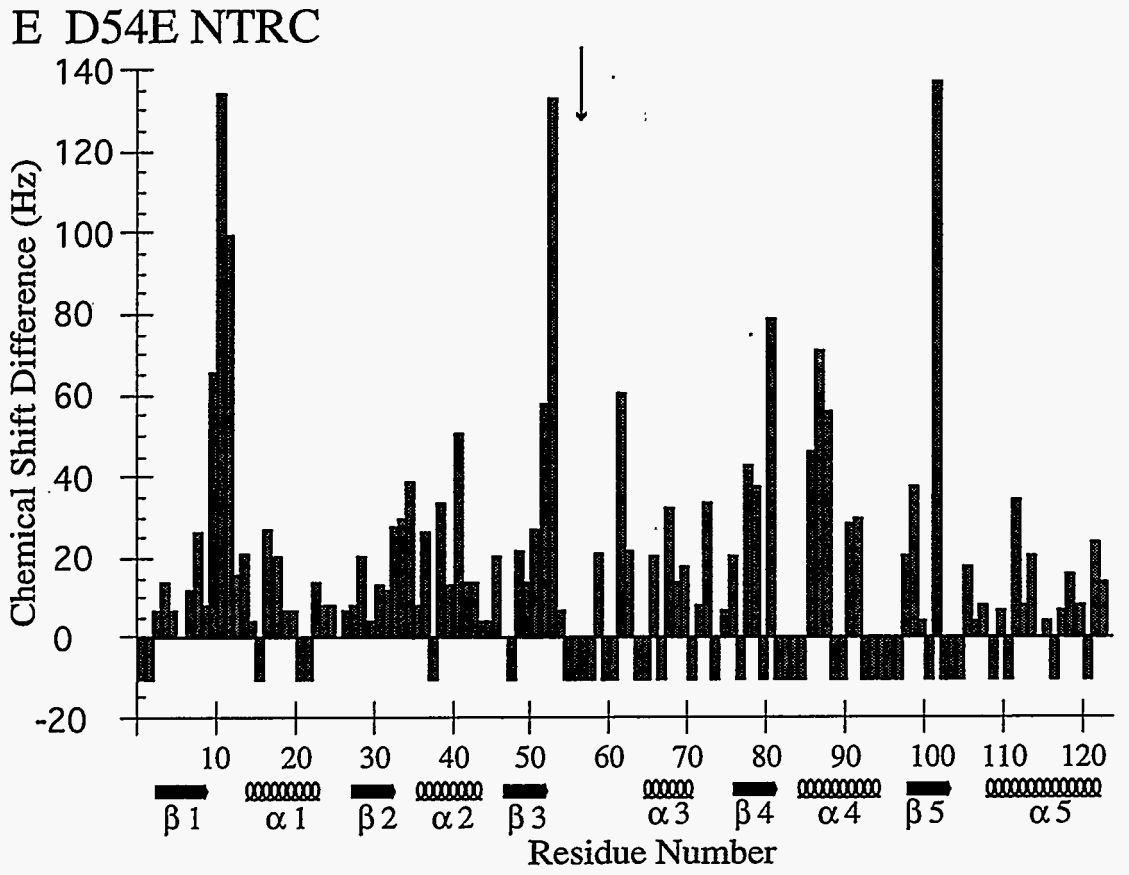


Figure 5.5: Graphs of the chemical shift differences between the constitutive mutants and wildtype NTRC plotted as a function of residue number. The negative bars represent residues for which either no amide was detectable ot the assignments for both forms could not be obtained. The arrows indicate the site(s) of mutation.

In order to provide a visual overview of the spatial distribution of the structural changes in the constitutive mutations, the residues in each mutant experiencing the largest chemical shift changes from wildtype NTRC are mapped onto the unphosphorylated wild type NTRC structure in Figures 5.6 - 5.10. The black spheres represent residues whose amides have a chemical shift changes of at least 50 Hz. This value was chosen as a compromise value to avoid visual clutter . The white spheres represent residues for which the chemical shift difference could not be determined.

Since the activity of the mutants increases as they are added in combination, these mutants were investigated with respect to a possible additivity of the shift changes. Figure 5.11 shows an overlay of the HSQC spectra of wildtype NTRC, the single mutant D86N NTRC, the double mutant D86N/A89T NTRC and the triple mutant D86N/A89T/V115I NTRC, enlarged to show only the peak arising from residue D11. The chemical shift change of D11 is consistent in direction for each of the mutants, with only the magnitude increasing with the number of mutations (which correlates to the activity). The other active site residues (D12 and D54) show similar behavior (data not shown).

Discussion

Chemical Shift Differences as a Probe of Conformational Change

In this study, chemical shift differences between active and inactive forms of NTRC are used as a probe of conformational change caused by activation. Chemical shift differences have long been used for mapping structural changes (Lowry et al., 1994; Stockman et al., 1995; Swanson et al., 1995). It must be noted that changes in chemical shifts are directly caused by changes in magnetic environments which must have a basis in conformational change. However, changes in chemical shifts are an indirect probe of conformational change and, therefore, the size of the change is not necessarily a direct reflection of the magnitude of change in conformation.

Chemical shift information for about 20% of the amides in NTRC could not be obtained in both the wild-type and the mutant form. For the most part, these represent

D86N NTRC

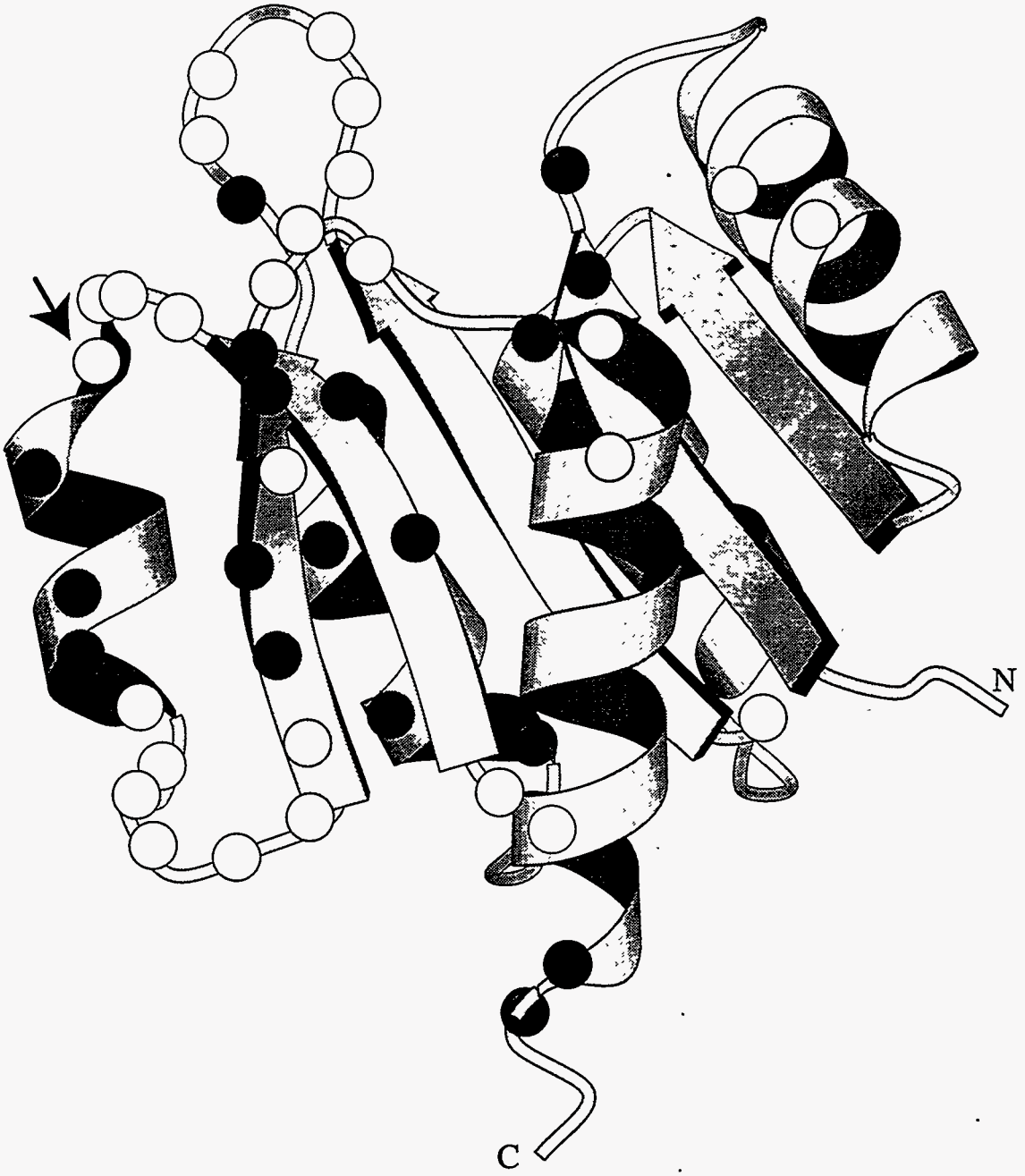


Figure 5.6: The largest chemical shift differences between D86N and wildtype NTRC mapped onto the structure of wildtype NTRC. The black spheres represent residues which have moved at least 50 Hz. The white spheres represent residues for which either amide was not detectable or the assignments for both D86N NTRC and wildtype NTRC were not obtained. The arrow indicates the site of mutation.

D86N/A89T NTRC

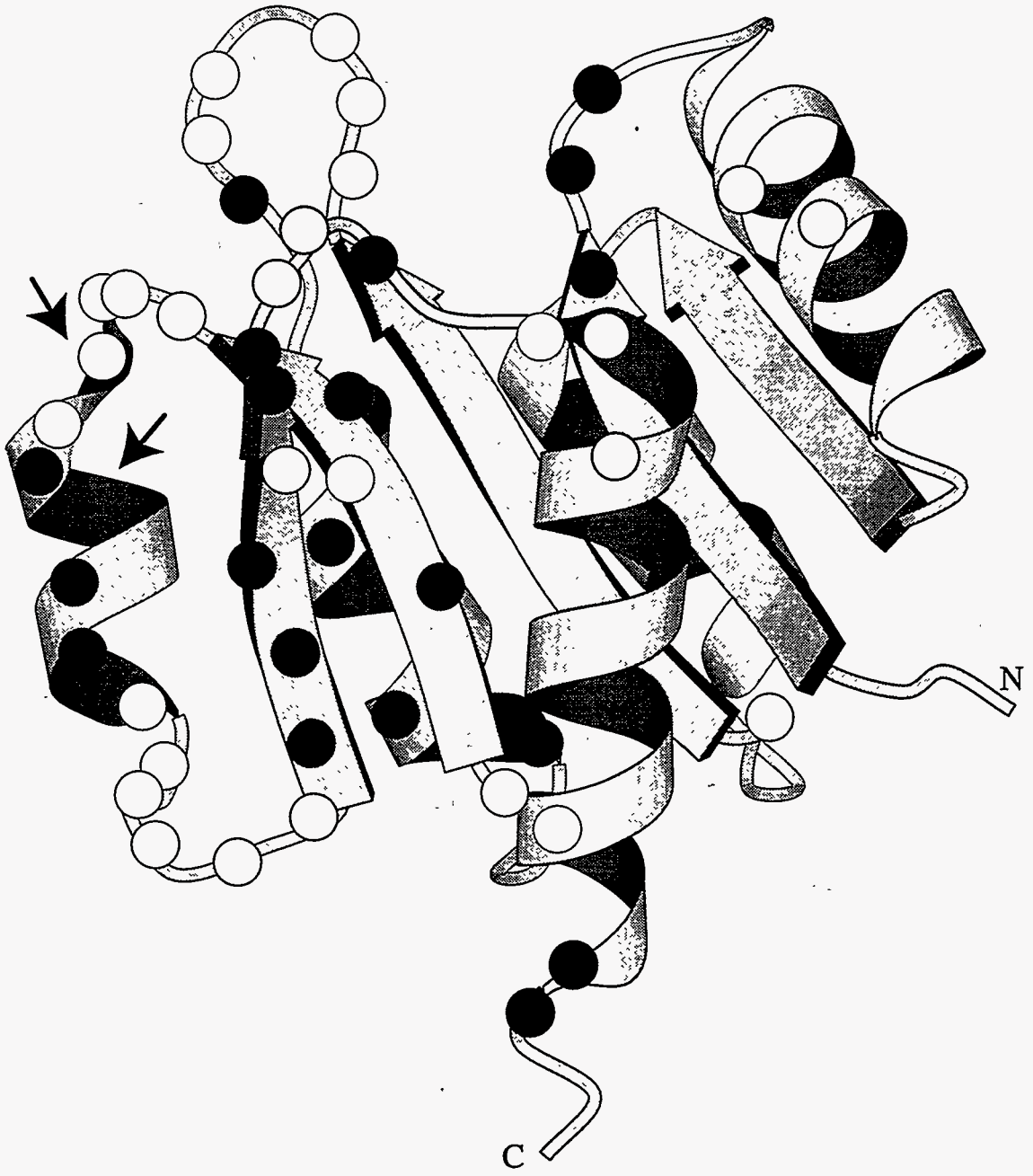


Figure 5.7: The largest chemical shift differences between D86N/A89T NTRC and wildtype NTRC mapped onto the structure of wildtype NTRC. The black spheres represent residues which have moved at least 50 Hz. The white spheres represent residues for which either the amide was not detectable or the assignments for both D86N/A89T and wildtype NTRC were not obtained. The arrows indicate the sites of mutation.

D86N/A89T/V115I NTRC

Figure 5.8: The largest chemical shift differences between D86N/A89T/V115I NTRC and wildtype NTRC mapped onto the the structure of wildtype NTRC. The black spheres represent residues which have moved at least 50 Hz. The white spheres represent residues for which either the amide was not detectable or both the assignmenns for both wildtype NTRC and D86N/A89T/V115I NTRC could not be obtained. The arrows indicate the sites of mutation.

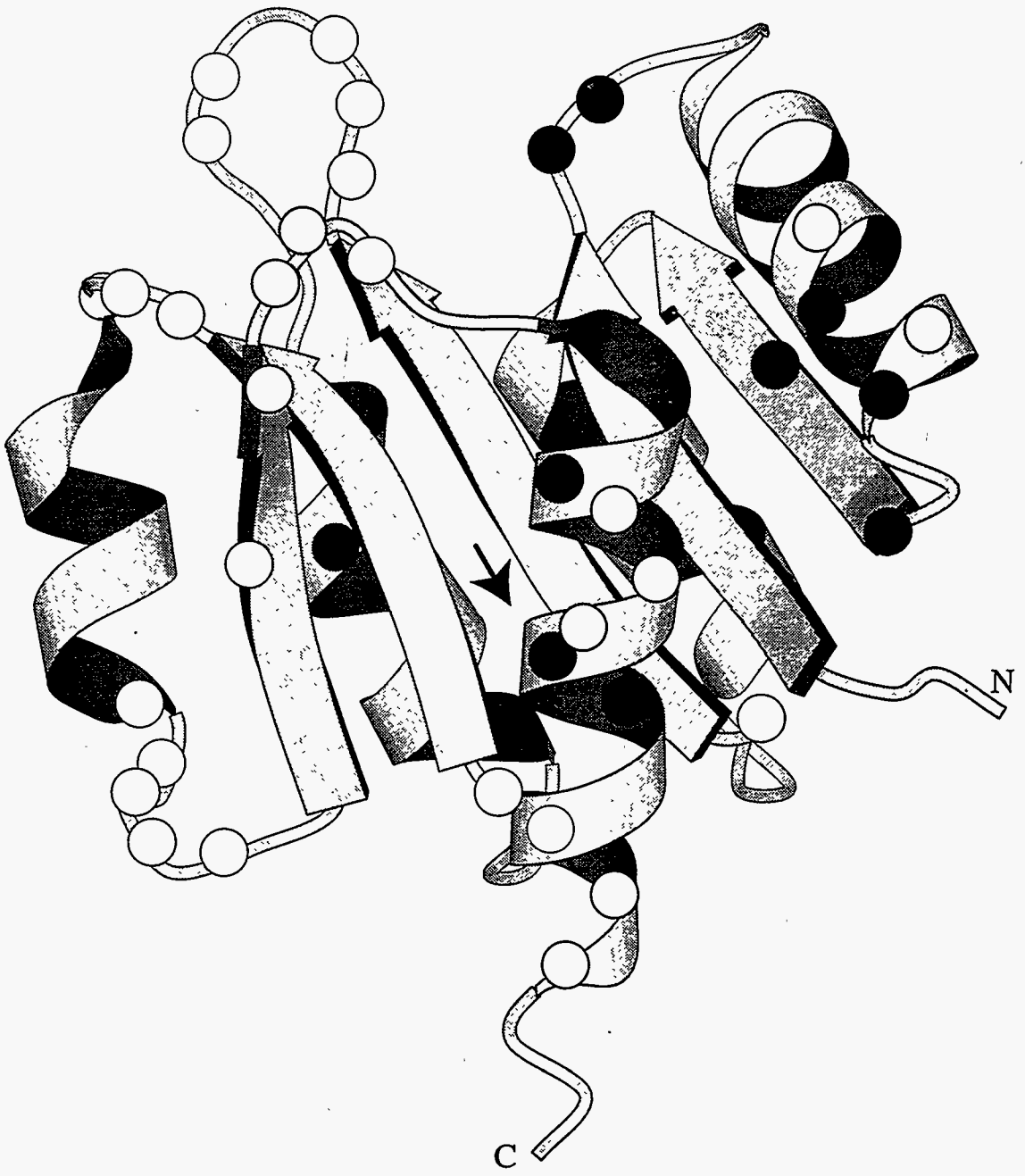


Figure 5.9: The largest chemical shift differences between V115INTRC and wildtype NTRC mapped onto a structure of wildtype NTRC. The black sheres represent residues which have moved at tleast 50 Hz. The white spheres represent residues for which either the amide was not detectable or both V115I and wildtype assignments were not obtained. The arrow indicates the site of mutation.

D54E NTRC

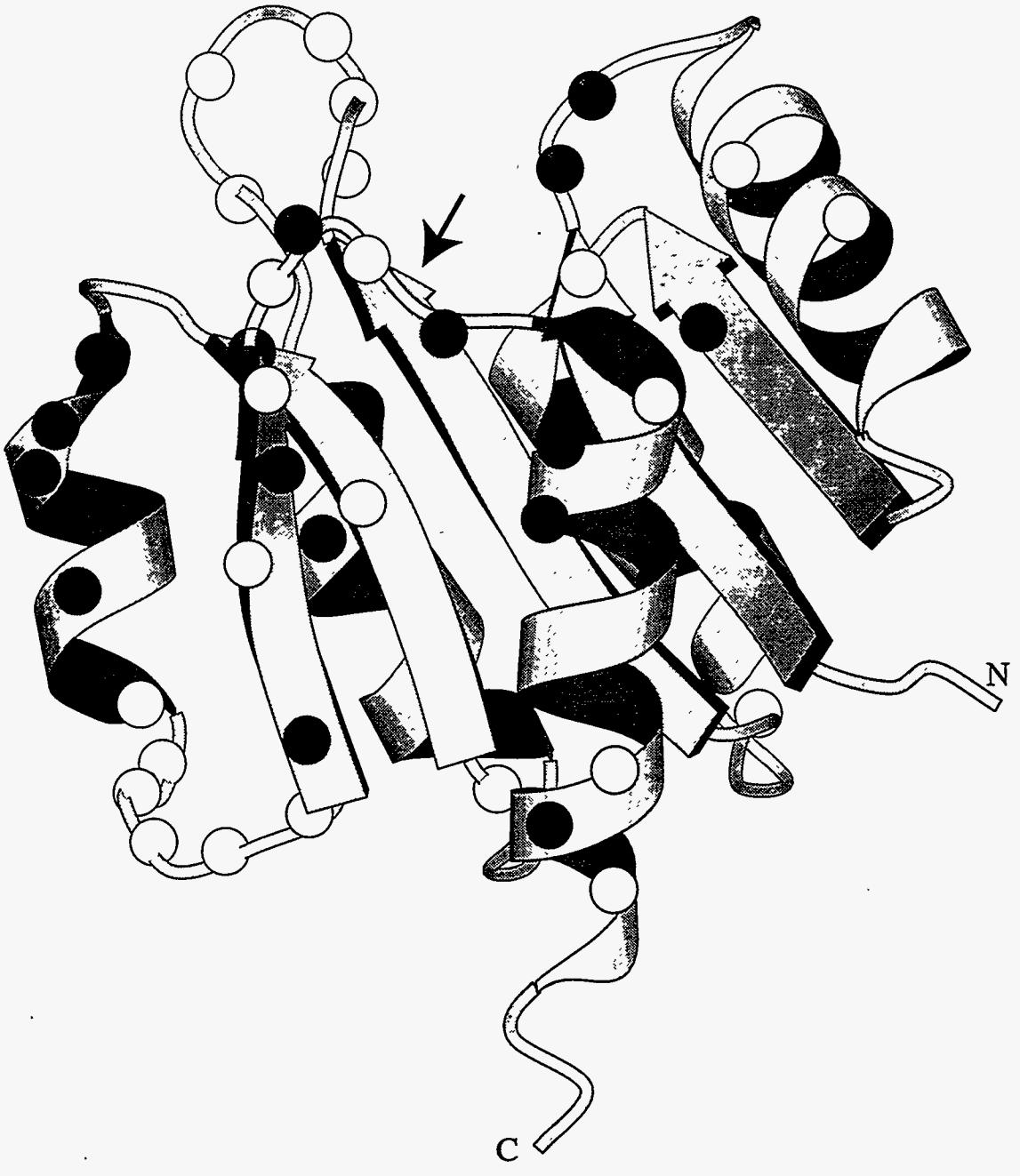


Figure 5.10: The largest chemical shift differences betwen D54E NTRC and wildtype NTRC mapped onto the structure of wildtype NTRC. The black spheres represent residues which have moved at least 50 Hz. The white spheres represent residues for which either the amide was not detectable or the assignments for both D54E NTRC and wildtype NTRC were not obtained. The arrow indicates the site of mutation.

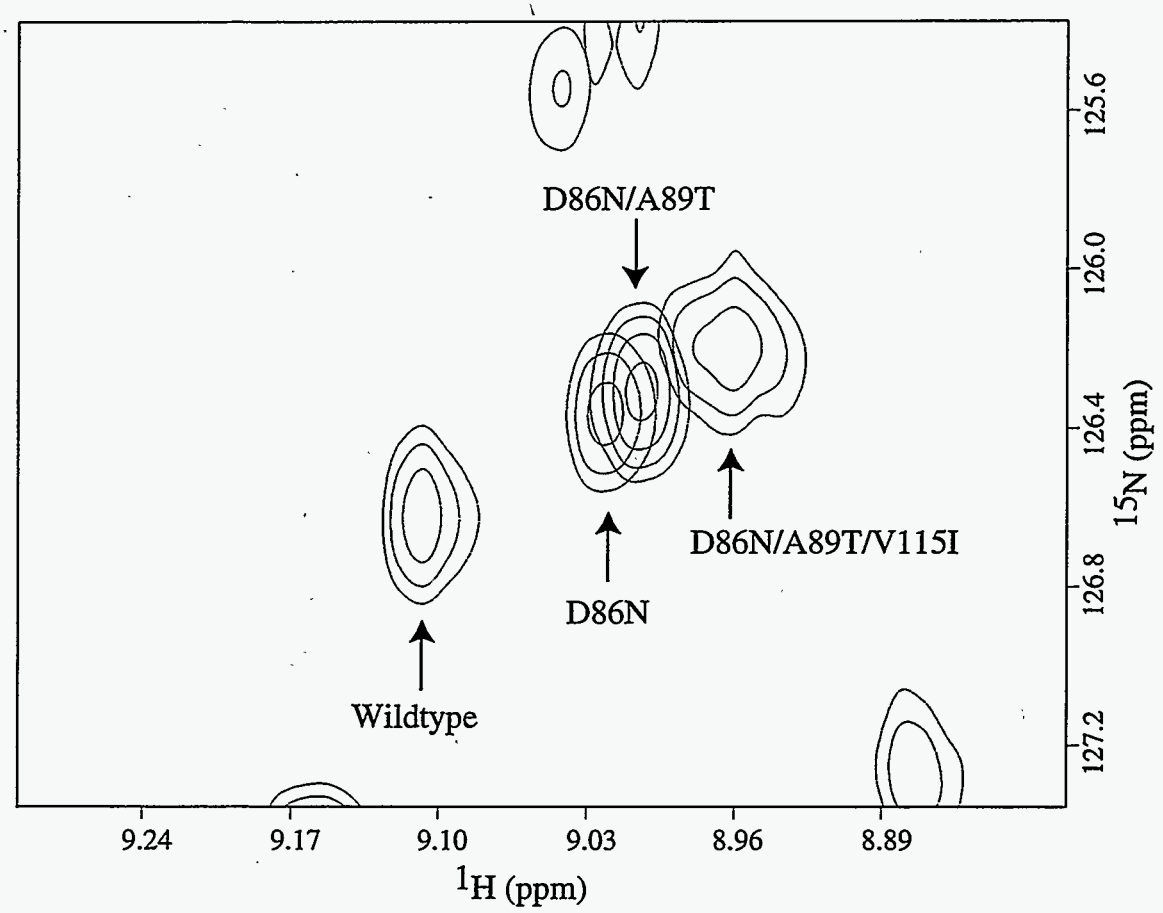


Figure 5.11: An overlay of ^{15}N-^{1}H HSQC spectra of residue D11 for wildtype, D86N, D86N/A89T and D86N/A89T/V115I NTRC.

residues for which the no amide signals could be detected because of high amide exchange rates. Thus, for instance, there are very few probes in this study for conformational change in the loop between helix 4 and strand 5. However, there are about 90 probes (out of a total possible of 118) which are well distributed throughout the molecule. These give a good overall picture of the spatial location of conformational change in the activated molecule.

Magnesium Binding and the Activation of NTRC

Mg^{2+} is required for the *in vitro* transcription assays that were used to characterize the mutants and is absolutely required for phosphorylation of wild type NTRC (see chapter 4). Thus it was of interest to ask whether Mg^{2+} is required for the activated state of NTRC. This was of particular importance for the NMR studies because high levels of Mg^{2+} cause precipitation of NTRC. The inability of D54E to bind Mg^{2+} while still being capable of activating transcription argues that magnesium binding is not required for activation of NTRC but is required for phosphorylation. This is in agreement with the data presented in chapter 4. This is in direct conflict with the model of receiver domain activation proposed of M. Coll and his colleagues which suggests that magnesium is the key factor for activation of the NTRC homologue CheY (Bellsolell et al., 1994) (see chapter 4).

The Active Form of NTRC

There are two possible structural models for the activation caused by the constitutive mutants. One model is that the mutated residues lie on a binding surface that interacts with the central domain and simply mimic the binding surface caused by the conformational change upon activation. By implication, there should only be local changes in conformation in these mutants around the site of mutation. The other model is that the inactive and active states of NTRC are in equilibrium, and the mutated residues serve to drive the equilibrium towards the activated state. This model would predict that the

mutations would result in large changes in conformation in regions distal to the site of mutation.

The present study confirms the second model. Constitutive mutations in helix 4 cause changes in conformation at the active site (D11, D12, D54, and T82) and the expected changes near the site of mutation as monitored by changes in chemical shift. Conversely, D54E, a constitutive mutant at the active site, causes changes in the conformation of the face of the molecule formed by helix 3, strand 4, helix 4 and strand 5, and the loop between strand 5 and helix 5 in addition to the expected changes in the active site. Thus, there seems to be a conformational link between these regions of NTRC.

The activity of the constitutive mutants when added in combination correlates well with the magnitude of the chemical shift changes at the active site. One explanation is that the constitutive mutations cause a partially activated structure to form and as the mutants are added in combination the structure becomes more and more like the fully active form of the molecule. The other possibility is that NTRC has two stable states - active and inactive. In this scenario, the constitutive mutations (or phosphorylation) cause the active state to become more populated. Neither of these ideas can be rigorously tested in the present study. However, the fact that the chemical shift changes are additive in the same direction upon the combination of constitutive mutations (Figure 5.11) is suggestive of the behavior of a residue undergoing fast exchange phenomena. In this model, the residues have different chemical shifts in the activated and nonactivated form. The position of the average signal between the two forms depend on the equilibrium constant. Addition of a second and third mutation shifts the equilibrium towards the activated form resulting in a shift of the average signal towards the signal of the fully activated form. Unfortunately, the location of the signal from D11 in the phosphorylated form of NTRC, which represents the fully active form, cannot be directly compared to the mutants, because the presence of Mg^{2+} causes extreme shifts in the active site.

V115I and Phosphodonor Specificity

V115I NTRC appears to be in a separate category from the rest of the constitutive mutants because it barely activates *in vitro* transcription. One possibility is that the mutation is involved in relaxing the phosphodonor specificity of NTRC. This would allow the mutant to phosphorylate itself from other sources in the absence of NTRB. These other sources might be small molecular phosphodonors such as acetyl phosphate and carbamyl phosphate or other histidine kinases (Fisher et al., 1995). This hypothesis is supported by the NMR data of the V115I mutant which shows significant chemical shift changes in the active site residues D11 and D12. However, the residues on the opposite side of the active site, T82 and L102, which makes contacts to the region formed by strand 4, helix 4 and strand 5, don't shift significantly. These are exactly the structural elements with the largest chemical shift changes in the other constitutive mutants. This suggests that conformational changes in T82 and/or L102 are required for activation of NTRC.

The Mechanism of Activation

The present study suggests that in addition to the active site itself, the face of the molecule formed by the loop between helix 3 and strand 4, helix 4, strand 5, and the loop between strand 5 and helix 5 undergoes a conformational change upon activation of NTRC. This is consistent with the region of the molecule that experiences the largest chemical shift changes upon phosphorylation (see chapter 4).

Furthermore, it suggests that this region of NTRC is conformationally linked to the active site. The residues T82 and L102 both shift significantly in all active mutants as well as in the phosphorylated wildtype NTRC. The residue L102 is of interest because it serves as a marker for the loop between strand 5 and helix 5 which has been speculated to be important for activation (Lowry et al., 1994). The amide of the active site residue K104 is not detectable in the HSQC probably due to fast amide exchange. However, L102 can be used as a probe for conformational changes on K104, since K104 is flanked by two highly conserved prolines making this loop more restricted since the angle ϕ is fixed in the proline rings. Consequently, conformational changes detected for L102 should reflect changes for

the whole loop. The constitutive mutations D86N and A89T both lie at the top of helix 4 very near T82 and this loop. The data support that T82 as well as the loop between strand 5 and helix 5 are the transducers for the signal from the active site to the "3445 face" of NTRC. These results would explain the lack of activity of the T87I mutant in CheY as well as the high conservation of an active site lysine followed by a *cis* proline in the loop between strand 5 and helix 5 in all members of the receiver domain family.

Chapter 6 Summary and Conclusions

The receiver domain superfamily provides a good system for the study of post-transitional control of regulatory proteins by phosphorylation. Considering the dearth of information about the structural basis of regulation by phosphorylation (Johnson, 1994), it was decided to pursue this question by studying the N-terminal receiver domain of NTRC. The phosphorylation of this molecule on D54 leads to activation of an ATPase in the central domain of the protein which leads, in turn, to transcriptional activation (Weiss et al., 1991).

Initially, the 3 dimensional structure of the N-terminal receiver domain of NTRC in the unphosphorylated state was determined by multidimensional heteronuclear NMR. This analysis demonstrated that the N-terminal domain of NTRC has an $(\alpha/\beta)5$ motif with a five stranded parallel β-sheet sandwiched by three α-helices on one side and two α-helices on the other. This is consistent with the 3 dimensional structure of CheY, a member of the receiver domain superfamily involved in chemotaxis, determined by X-ray crystallography (Volz & Matsumura, 1991).

This structure alone does not provide insight into the mechanism of control of NTRC by phosphorylation. A comparison of this structure with a phosphorylated structure of NTRC could reveal a great deal about the activation process. Unfortunately, the short half-life of the phosphorylated form of NTRC makes these investigations extremely difficult. At the present time, conditions have been obtained which should allow the assignment of the backbone and sidechain resonances of the phosphorylated form of NTRC. It is not clear whether these conditions are sufficient for the determination of the 3 dimensional structure of this form.

Tentative chemical shift differences between unphosphorylated and phosphorylated forms of NTRC can be used as a probe of the regions of the molecule undergoing conformational rearrangement. This type of analysis suggests that phosphorylation causes

structural changes in the loop between strand 1 and helix 2, the loop between strand 2 and helix 2, helix 3, strand 4, helix 4, strand 5 and the loop between strand 5 and helix 5. Since the aforementioned loops form the active site, it is unsurprising that these areas should be rearranged upon phosphorylation. The other perturbed region, the face of the molecule formed by helix 3, strand 4 , helix 4, and strand 5 (henceforth referred to as the "3445 face" of NTRC) would seem to be a good candidate for the part of the molecule that transduces the signal to the central domain.

Two other lines of inquiry direct attention to this region of NTRC. First, dynamic analysis of the backbone of NTRC shows that helix 4 and strand 5 experience an unusually high degree of dynamic motion for secondary structural elements. Second, two mutants of NTRC, D86N and A89T, constitutive for activation lie in helix 4 (Flashner et al., 1995).

Further investigation of the consitutive mutants, alone and in combination, suggests that the active site and the "3445 face" are conformationally linked. That is, activating mutations in the active site lead to changes in the "3445 face" and vice versa while the rest of the molecule is relatively undisturbed. This linkage is probably mediated by the active site residues T82 and K104 which lie close to the top of helix 4. The fact that the constitutive mutants and the phosphorylated wildtype form of NTRC experience conformational changes in similar regions suggests that the 3 dimensional structure of a constitutive mutant will shed light on the basis of activation.

These data taken together lead me to propose the following model. The "3445 face" of NTRC is packed against the central domain of NTRC in the full length protein (hence the loss of these contacts leads increased in dynamics in this region in the N-terminal construct). Upon phosphorylation, electrostatic interactions lead to a rearrangement of the active site residues T82 and K104. This, in turn, leads to a rearrangement of the "3445 face" of NTRC which transduces a the signal to the central domain.

In order to test this model and to provide a more detailed picture of the structural basis of the regulation of NTRC by phosphorylation, more information is needed.

Specifically, a structure of the phosphorylated form of NTRC (or a constitutive mutant) should allow some aspects of the model to be tested. However, other types of experiments involving the central domain will be required to fully explain the molecular basis of phosphorylation control of NTRC.

References

Abragam, A. (1961) *Principles of Nuclear Magnetism*, Clarendon Press, Oxford.

Akke, M., Carr, P.A. & Palmer, A.G. (1994)*Journal of Magnetic Resonance Series B 104*, 298-302.

Aue, W.P., Bartholdi, E. & Ernst, R.R. (1976)*Journal of Chemical Physics 64*, 2229-2246.

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

Babu, Y.S., Buggs, C.E. & Cook, W.J. (1988)*Journal of Molecular Biology 204*, 191-204.

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

Barford, D. & Johnson, L.N. (1989)*Nature 340*, 609-66.

Bax, A. (1994)*Curr. Opin. Struct. Biol. 4*, 738-744.

Bax, A., Clore, G.M. & Gronenborn, A.M. (1990)*Journal of Magnetic Resonance 88*, 425-431.

Bax, A. & Davis, G. (1985)*Journal of Magnetic Resonance 65*, 355-360.

Bax, A. & Grzesiek, S. (1993)*Accounts in Chemical Research 26*, 131-138.

Bax, A. & Pochapsky, S.S. (1992)*Journal of Magnetic Resonance 99*, 638-643.

Bellsolell, L., Prieto, J., Serrano, L. & Coll, M. (1994)*Journal of Molecular Biology 238*, 489-495.

Blum, K. (1981) *Density Matrix Theory and Its Applications*, Plenum Press, New York.

Bodenhausen, G. & Ruben, D.J. (1980)*Chemical Physics Letters 69*, 185-189.

Bourret, R.B., Borkovich, K.A. & Simon, M.I. (1991)*Ann. Rev. Biochem. 60*, 401-441.

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. (1983a)*Journal of Magnetic Resonance 53*, 521-528.

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

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

Chock, P.B., Shacter, E., Jurgensen, S.R. & Rhee, S.G. (1985)*Current Topics in Cellular Regulation 27*, 3-12.

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.

Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C. & Gronenborn, A.M. (1990)*Journal of the American Chemical Society 112*, 4989-4991.

Colloc'h, N. & Cohen, F.E. (1991)*Journal of Molecular Biology 221*, 603-613.

Dean, A.M. & D.E., Koshland. (1990)*Scince 249*, 1044-46.

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

Drake, S.K. (1993)*Journal of Biological Chemistry 268*, 13089-13096.

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

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

Erikson, M., Hard, T. & Nilsson, L. (1995)*Biophysical Journal 69*, 329-339.

Ernst, R.R. & Anderson, W.A. (1966)*Review of Scentific Instruments 37*, 93-100.

Fano, U. (1957)*Reviews of Modern Physics 29*, 74-103.

Farrar, T.C. & Harriman, J.E. (1995) *Density Matrix Theory and Its applications to NMR Spectroscopy.*, The Farragut Press, Madison.

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

Fesik, S.W., Eaton, H.L., Olejniczak, E.T. & Zuiderweg, E.R.P. (1990)*Journal of the American Chemical Society 112*, 886-888.

Fiedler, U. & Weiss, V. (1995)*EMBO 14*, 3696-3705.

Fisher, S.L., Jiang, W., Wanner, B.L. & Walh, C.T. (1995)*Journal of Biological Chemistry 270*, 23143-9.

Flashner, Y., Weiss, D.S., Keener, J. & Kustu, S. (1995)*Journal of Molecular Biology 249*, 700-13.

Ganguli, S., Wang, H., Matsumura, P. & Volz, K. (1995)*Journal of Biological Chemistry 270*, 17386-17393.

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

Goldman, M. (1988) *Quantum Description of High Resolution NMR in Liquids*, Clarendon Press, New York.

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

Gronenborn, A.M. & Clore, G.M. (1994)*Journal of Biomolecular NMR 4*, 455-458.

Grzesiek, S. & Bax, A. (1992a)*Journal of the American Chemical Society 114*, 6291-6293.

Grzesiek, S. & Bax, A. (1992b)*Journal of Magnetic Resonance 99*, 201-207.

Grzesiek, S., Vuister, G.W. & Bax, A. (1993)*Journal of Biomolecular NMR 3*, 487-93.

Guntert, P., Braun, W., Billeter, M. & Wüthrich, K. (1989)*Journal of the American Chemical Society 111*, 3997-4004.

Hardie, D. (1993) *Protein Phosphorylation: A Practical Approach*, Oxford University Press, Oxford.

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

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

Heidorn, D.B. & Trewhella, J. (1988)*Biochemistry 27*, 909-915.

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.

Hirschman, J., Wong, P.K., Sei, K., Keener, J. & Kustu, S. (1985)*Proceedings of the National Academy of Sciences USA 82*, 7525-9.

Howarth, M.A., Lain, L.Y., Hawkes, G.E. & Sales, K.D. (1986)*Journal of Magnetic Resonance 68*, 433-452.

Hubbard, J.S. & Stadtman, E.R. (1967)*Journal of Bacteriology 93*, 1045-55.

Hunt, T.P. & Magasanik, B. (1985)*Proceedings of the National Academy of Sciences USA 82*, 8453-7.

Hunter, T. & Sefton, B.M. (1991) *Protein Phosphorylation*, Academic Press, San Diego.

Hurley, J.H., Dean, A.M., Sohl, J.L. & Koshland, D.E. (1990)*Science 249*, 1012-16.

Ikura, M., Kay, L.E. & Bax, A. (1990)*Biochemistry 29*, 4659-4667.

Jeener, J. (1971) Ampere Summer School, Basko Polje, Yugoslavia

Jeener, J., Meier, B.H., Bachmann, P. & Ernst, R.R. (1979)*Journal of Chemical Physics 71*, 4546-4553.

Johnson, L.N. (1994)*Nature Struct. Biol. 1*, 657-9.

Johnson, L.N. & Barford, D. (1993)*Annual Reveiw of Biophysics and Biomolecular Structure 22*, 199-232.

Kay, L.E. (1995)*Curr. Opin. Struct. Biol. 5*, 674-681.

Kay, L.E. & Bax, A. (1990)*Journal of Magnetic Resonance 86*, 110-126.

Kay, L.E., Marion, D. & Bax, A. (1989)*Journal of Magnetic Resonance 84*, 72-84.

Kay, L.E., Muhandiram, D.R., Farrow, N.A., Aubin, Y. & Forman-Kay, J. (1996) *Biochemistry 35*, 361-368.

Kay, L.E., Torchia, D.A. & Bax, A. (1989)*Biochemistry 28*, 8972-8979.

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

Klose, K., North, A., Stedman, K. & Kustu, S. (1994)*Journal of Molecular Biology 240*, 234-242

Klose, K.E., Weiss, D.W. & Kustu, S. (1993)*Journal of Molecular Biology 232*, 67-78.

Lipari, G. & Szabo, A. (1982a)*Journal of the American Chemical Society 104*, 4559-4570.

Lipari, G. & Szabo, A. (1982b)*Journal of teh American Chemical Society 104*, 4546-4559.

Live, D.H., Davis, D.G., Agosta, W.C. & Cowburn, D. (1984)*Journal of the American Chemical Society 106*, 1934-1941.

Lowry, D.F., Roth, A.F., Rupert, P.B., Dahlquist, F.W., Moy, F.J., Domaille, P.J. & Matsumura, P. (1994)*The Journal of Biological Chemistry 269*, 26358-26362.

Lukat, G.S., Lee, B.H., Mottonen, J.M., Stock, A.M. & Stock, J.B. (1991)*Journal of Biological Chemistry 266*, 8348-8354.

Lukat, G.S., McCleary, W.R., Stock, A.M. & Stock, J.B. (1992)*Proceedings of the National Academy of Science USA 89*, 718-722.

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

Macura, S., Huang, Y. & R.R., Ernst. (1981)*Journal of Magnetic Resosnace 43*, 259-281.

Mandel, A.M., Akke, M. & Palmer, A.G. (1995)*Journal of Molecular Biology 246*, 144-163.

Manoleras, N. & Norton, R.S. (1992)*Journal of Biomolecular NMR 2*, 485-494.

Marion, D., Ikura, M., Tschudin, R. & Bax, A. (1989a)*Journal of Magnetic Resonance 85*, 393-399.

Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. & Bax, A. (1989b)*Journal of the American Chemical Society 111*, 1515-1517.

Mateescu, G.D. & Valeriu, A. (1993) *2D NMR: Density Matrix and Product Operator Treatment*, P T R Prentice Hall, Englewood, N.J.

Matsushima, N., Izumi, Y., Matsuo, T., Yoshino, H., Ueki, T. & Miyake, Y. (1989) *Journal of Biochemistry (Tokyo) 105*, 883-887.

McConnell, J. (1987) *The Theory of Nuclear Magnetic Relaxation in Liquids*, Cambridge University Press, New York.

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

Merrick, M.J. & Edwards, R.A. (1995)*Microbiological Reviews 59*, 604-22.

Mettke, I., Fiedler, U. & Weiss, V. (1995)*Journal of Bacteriology 177*, 5056-61.

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

Morris, G.A. & Freeman, R. (1979)*Journal of the American Chemical Society 101*, 760.

Mueller, L. (1979)*Journal of the American Chemical Society 101*, 4481-4484.

Munowitz, M. (1988) *Coherence and NMR*, John Wiley and Sons, New York.

Neri, D., Szyperski, T., Otting, G., Senn, H. & Wüthrich, K. (1989)*Biochemistry 28*, 7510-7516.

Neuhaus, D. & Williamson, M. (1989) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH Publishers, New York.

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.

Nixon, B.T., Ronson, C.W. & Ausubel, F.M. (1986)*Proceedings of the National Academy of Sciences of the United States 83*, 7850-4.

Orekhov, V.Y., Nolde, D.E., Golovanov, A.P. & Korzhnev, D.M. (1995)*Applied Magnetic Resonance 9*, 581-588.

Otting, G. & Wutrich, K. (1988)*Journal of Magnetic Resonance* 569-574.

Packer, K.J. & Wright, K.M. (1983)*Molecular Physics 50*, 797-813.

Palmer, A.G. (1993)*Curr. Opin. Biotech. 4*, 385-391.

Palmer, A.G., Cavanagh, J., Wright, P.E. & Rance, M. (1991)*Journal of Magnetic Resonance 93*, 151-170.

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

Popham, D.L., Szeto, D., Keener, J. & Kustu, S. (1989)*Science 243*, 629-635.

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

Quon, K.C., Marczynski, G.T. & Shapiro, L. (1996)*Cell 84*, 83-93.

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

Redfield, C., Boyd, J., Smith, L.J., Smith, R.A.G. & Dobson, C.M. (1992)*Biochemistry 31*, 10431-37.

Reitzer, L.J. & Magasanik, B. (1985)*Proc. Natl. Acad. Sci. USA 82*, 1979-1983.

Rhee, S.G., Chock, P.B. & Stadtman, E.R. (1989)*Advances in Enzymology and Related Areas of Molecular Biology 62*, 37-92.

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

Roman, S.J., Meyers, M., Volz, K. & Matsumura, P. (1992)*Journal of Bacteriology 174*, 6247-55.

Ronson, C.W., Nixon, B.T. & Ausubel, F.M. (1987)*Cell 49*, 579-81.

Sakamoto, N., Kotre, A.M. & Savageau, M.A. (1975)*Journal of Bacteriology 124*, 775-83.

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)*Journal of Bacteriology 174*, 5112-5122.

Saunders, M., Wishnia, A. & Kirkwood, J.G. (1957)*Journal of the American Chemical Society 79*, 3289-93.

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.

Seaton, B.A., Head, J.F., Engleman, D.M. & Richards, F.M. (1985)*Biochemistry 24*, 6740-6743.

Sheridan, R.C., McCullough, J.F. & Wakefield, Z.T. (1971)*Inorganic Synthesis 13*, 23-26.

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

Skelton, N.J., Palmer, A.G., Akke, M., Kordel, J., Rance, M. & Chazin, W. (1993a)*Journal of Magnetic Resonance, Series B 102*, 253-264.

Sklenar, V. & Bax, A. (1987)*Journal of Magnetic Resonance 74*, 469-479.

Sockett, H., Yamaguchi, S., Kihara, M., Irikura, V.M. & Macnab, R.M. (1992)*Journal of Bacteriology 174*, 793-806.

Son, H.S. & Rhee, S.G. (1987) *Journal of Biological Chemistry 262*, 8690-5.

Sorenson, O.W., Eich, G.W., Levitt, M.H., Bodenhausen, G. & Ernst, R.R. (1983)*Progress in NMR Spectrscopy 16*, 163-92.

Spera, S. & Bax, A. (1991)*Journal of the American Chemical Society 113*, 5490-5492.

Sprang, S.R., Acharya, K.R., Goldsmith, E.J., Stuart, D.I., Varvill, K., R.J., F., Madsen, N.B. & Johnson, L.N. (1988)*Nature 336*, 215-21.

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. (1989a)*Nature 337*, 745-749.

Stock, J.B., Ninfa, A.J. & Stock, A.M. (1989b)*Microbiological Reviews 53*, 450-490.

Stockman, B.J., Bannow, C.A., Miceli, R.M., Degraaf, M.E., Fischer, H.D. & Smith, C.W. (1995)*International Journal of Peptide and Protein Research 45*, 11-16.

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

Swanson, R.V., Lowry, D.F., Matsumura, P., McEvoy, M.M., Simon, M.I. & Dahlquist, F.W. (1995) *Nature Struct. Biol. 2*, 906-910.

van de Ven, F.J.M. (1995) *Multidimensional NMR in Liquids*, VCH publishers, New York.

Van De Ven, F.J.M. & Hilbers, C.W. (1983)*Journal of Magnetic Resonance 54*, 512-520.

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

Volz, K. & Matsumura, P. (1991)*Journal of Biological Chemistry 296*, 15511-15519.

Vuister, G., Wang, A.C. & Bax, A. (1993)*Journl of the AmericanChemial Society 115*, 5334-35.

Vuister, G.W. & Bax, A. (1992)*Journal of Magnetic Resonance 98*, 428-435.

Wagner, G. (1993)*Curr. Opin. Struct. Biol. 3*, 748-753.

Wanner, B.L. (1992)*Journal of Bacteriology 174*, 2053-58.

Wedel, A. & Kustu, S. (1995)*Genes Develop. 9*, 2042-52.

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

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)*Journal of Molecular Biology 182*, 295-315.

Wishart, D.S., Sykes, B.D. & Richards, F.M. (1991)*Journal of Molecular Biology 222*, 311-333.

Woolfolk, C.A., Shapiro, B. & Stadtman (1966) *Archives of Biochemistry and Biophysics 116*, 177-92.

Woolfolk, C.A. & Stadtman, E.R. (1967)*Archives of Biochemistry and Biophysics 118*, 736-55.

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)*Journal of Molecular Biology 169*, 949-961.

Yamazaki, T., Lee, W., Arrowsmith, C.H., Muhandiram, D.R. & Kay, L.E. (1994)*Journal of the American Chemical Society 116*, 11655-11666.

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