Solid state and solution NMR studies of model systems for parts of membrane proteins
by Earle Raymond Adams

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry
Montana State University

Abstract:
Detailed atomic scale structural information on proteins and other biomolecules is needed to understand biological mechanisms. However, many biologically important macromolecules such as large soluble proteins or membrane proteins normally cannot be studied directly by powerful new methods of liquid state NMR due to broad lines inherent in slowly tumbling large molecules. Paramagnetic shift reagents were investigated for use in solid state NMR of macromolecules and EDTA was used as a model for detecting sidechain carboxyls that bind metal ions in proteins. The ς22 tensor from the chemical shift anisotropy (CSA) of the carbonyl carbons was used to distinguish the OCO angle in metal ion complexes in the solid state. Liquid state 2D NMR was used to study a potential metal ion binding site in bacteriorhodopsin (BR); A peptide fragment that is homologous to an EF-hand metal ion binding loop between the C&D helices in BR. This peptide bound La+3 and showed secondary structure in solution which fit into the BR helix structure and the EF-hand coordination could be completed by 161Glu on the EF loop of BR. Protein-protein interactions, important in biological processes, are difficult to study by solution NMR methods because of the high molecular weight of the complexes. We used Tr-NOESY to study the protein bound conformation of a peptide ligand in exchange with a receptor. The p47-phox bound structure of a carboxyl terminal peptide from gp91-phox was studied since this peptide blocks superoxide generation by binding to the p47-phox control protein. The gp91-phox carboxyl terminal peptide is in fast exchange with p47-phox relative to cross relaxation and binds to p47-phox in an extended open loop conformation involving strong interaction of certain peptide sidechains.

SOLID STATE AND SOLUTION NMR STUDIES OF MODEL SYSTEMS FOR PARTS OF MEMBRANE PROTEINS

by

Earle Raymond Adams

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Chemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 1994

APPROVAL

of a thesis submitted by

Earle Raymond Adams

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

8/4/94
Date

Edward A. Hoat
Chairperson, Graduate Committee

Approved for the Major Department

8/5/94
Date

David M Dooley
Head, Major Department

Approved for the College of Graduate Studies

8/8/94
Date

R L Brown
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation for sale in and from microform or electronic format, along with the right to reproduce and distribute my abstract in any format in whole or in part."

Signature Earle A. Adams

Date 8/8/94

ACKNOWLEDGEMENTS

I would like to thank the people who have been important in my life and fundamental in finishing my work at Montana State University. People in my research group who have been a great source of help: Craig Johnson for synthesizing peptides needed for my work, Laura Holte for collaboration in solid state NMR work, Julie Furstenau for close collaboration and discussion of NMR and structure work, Chris Lambert for his great aid in computer macro design for NMR processing and teaching me the use of computers in my research, David Poole for his aid in programming macros needed for NMR refinement work, Sam Helgerson for his help in introducing me to solid state NMR.

I would also like to acknowledge some special colleagues who have been extremely helpful and made research and life enjoyable: Steve Dunham for his constant help in NMR experiments and applications, Brian Williams who has been a great friend and companion over the years and has played a role with me on a range of activities from conversing about NMR experiments to fantastic fishing expeditions, Rajani Bommakanti who has been around for support and friendship since my start, Lee David for his constant help through the years in trouble shooting NMR problems, Mark Quinn for his support and collaboration on a major part of my research.

I would like to make a very special thanks to my parents Mr. and Mrs. Earle G. Adams who have made all of this possible, who have always given their support and faith in my life, education and making this endeavor possible. I would also like to thank my wife Angie Adams for her patience and support.

A special thanks to my advisor Edward Dratz for his support and help over my research years at Montana State University, and for allowing variability in my research to follow different paths.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Detailed atomic scale structural information on proteins and other biomolecules is needed to understand biological mechanisms. However, many biologically important macromolecules such as large soluble proteins or membrane proteins normally cannot be studied directly by powerful new methods of liquid state NMR due to broad lines inherent in slowly tumbling large molecules. Paramagnetic shift reagents were investigated for use in solid state NMR of macromolecules and EDTA was used as a model for detecting sidechain carboxyls that bind metal ions in proteins. The σ_{22} tensor from the chemical shift anisotropy (CSA) of the carbonyl carbons was used to distinguish the OCO angle in metal ion complexes in the solid state. Liquid state 2D NMR was used to study a potential metal ion binding site in bacteriorhodopsin (BR); A peptide fragment that is homologous to an EF-hand metal ion binding loop between the C&D helices in BR. This peptide bound La^{+3} and showed secondary structure in solution which fit into the BR helix structure and the EF-hand coordination could be completed by 161Glu on the EF loop of BR. Protein-protein interactions, important in biological processes, are difficult to study by solution NMR methods because of the high molecular weight of the complexes. We used Tr-NOESY to study the protein bound conformation of a peptide ligand in exchange with a receptor. The p47-*phox* bound structure of a carboxyl terminal peptide from gp91-*phox* was studied since this peptide blocks superoxide generation by binding to the p47-*phox* control protein. The gp91-*phox* carboxyl terminal peptide is in fast exchange with p47-*phox* relative to cross relaxation and binds to p47-*phox* in an extended open loop conformation involving strong interaction of certain peptide sidechains.

CHAPTER 1

INTRODUCTION

Great advances in understanding biological mechanisms have been made when detailed atomic scale structural information on proteins and other biomolecules can be obtained. Historically, x-ray crystallography has provided the principal technique for revealing biomolecular structure. In the past ten years or so NMR and to a lesser extent electron cryomicroscopy have come to be increasingly useful techniques for macromolecular structure determination.

NMR has the distinction, of the three structural techniques mentioned, of being able to study soluble proteins in close to their natural environments. X-ray crystallography while a powerful structure technique is seriously limited by the ability to prepare well ordered three dimensional crystals and electron cyromicroscopy is even more limited by the ability to prepare planar sheets of well ordered 2D crystals.

For protein structure determination by NMR, chemically shifted resonances for all the nuclei of interest in the molecule must be assigned. NOE (nuclear overhauser effect) and coupling constant data which can be gathered from nuclei such as protons can then be used to provide interatomic

distance and dihedral angle information respectively. A major obstacle in NMR is the assignment of proton resonances in larger molecules due to the spectral complexity and broader lines at higher molecular weights.

In x-ray crystallography the electron density of an object is defined by the sum of sine waves described by frequency, amplitude, and phase. The frequencies are determined by the lattice spacing of the molecules in the crystals. The phase contributes position information. Diffraction intensities contribute the amplitude (actually the squares of the amplitudes) which give the main contribution to the electron densities. Since amplitude and phase are not measured directly, the major task in x-ray diffraction is phase determination using such methods as isomorphous replacement, molecular replacement, and anomalous scattering (Rhodes, 1993).

Electron microscopy has a theoretical advantage over x-ray in that amplitude and phase information can be obtained by Fourier transformation of an image. The downside of electron microscopy is that images are imperfect and are subject to instrument noise, sample movement, and other shortcomings. Amplitude information can be reconstructed better from crystal diffraction than from electron microscopy images. However phase information is recovered efficiently in electron microscopy through various computational procedures such as Fourier and back projection

methods (Chiu, 1993).

Sample preparation is relatively straightforward for NMR if the solubility and conformational stability of the molecules of interest is not a problem. X-ray relies on well ordered 3D crystals which these days are usually fast frozen, while 2D sheet samples are observed in vitreous ice or glucose in electron microscopy typically at a temperature of -150° C (or lower). Most biological specimens do not appear as crystalline arrays naturally in vivo. Crystallization procedures are not straightforward and electron beams used in the electron microscopy experimental procedure tend to damage the sample. Due to the difficulty of preparation of well ordered 2D crystalline arrays in biospecimens and instrument limitations, electron microscopy resolution in the very best case at this time is less than 3.5 Å. Therefore electron microscopy is used to a much lesser extent than x-ray crystallography and NMR.

Through x-ray diffraction analysis one average structure is normally obtained, but such is not the case in NMR. An ensemble of structures is typically obtained from NMR in which structures differ to a greater or lesser extent among themselves, by their different local atom positions reflecting motions and by limitations in defining structure due to limited data (Kline, et al., 1986). One potential advantage of NMR, with the use of relaxation times, is the capability of studying the dynamics of molecular motion.

There are an increasing number of cases where the combined use of x-ray crystallography and NMR are being used to study the structure of proteins. One such case is the structure of the elastase inhibitor elgin C. NMR was done on the free inhibitor while the X-ray was done on complexes with proteases (subtilsin, and thermistase). The calculated NMR structures had an RMSD agreement of the alpha-carbon positions of 0.61 Å for all the structures. The x-ray structure had a resolution of 0.455 Å for the alpha-carbons. Comparison of the backbone atoms between the x-ray structure and NMR structures mainly showed a difference in the proteinase binding loop which was altered by interactions in the complex. This study (and others) showed that the NMR structures of small proteins can be determined to almost an equivalent precision/accuracy of a good x-ray structure (Billeter et al., 1989).

It has been seen in several comparisons between NMR and x-ray structures that NMR can be as precise as x-ray in the interior of a protein. X-ray and NMR structure comparisons differ mainly on protein surfaces. X-ray structures determined in different crystal forms may show larger differences in surface loop conformations than are found in an ensemble of NMR structures. This difference in the surface residues between x-ray and NMR structures of proteins is probably due to intermolecular contacts in the crystal which may distort surface residues of proteins.

Intermolecular contacts in crystals and the mobility of surface residues gives NMR an advantage of more accurately determining the structure and dynamics of binding loops on a proteins's surface over x-ray crystallography and electron microscopy (Wagner et al., 1992).

Two-dimensional NMR (2D-NMR) is a powerful NMR technique that has been in use for over ten years (Wuthrich, 1986). Two dimensional NMR is an extension of 1D NMR with the advantage of increasing resolution by spreading the chemical shifts of resonances over two axes. Two dimensional NMR has been sucessfully used to study the molecular structure of smaller biomolecules up to about 10,000 dalton molecular weight. The use of 2D NMR of biomolecules of over 10,000 daltons becomes more difficult due to large line widths and spectral complexity. Due to the limitations of 2D NMR, higher dimension NMR experiments have been developed as tools for spectral and structure elucidation of higher molecular weight biomolecules (Clore et al., 1991). Three and 4D NMR of proteins require full replacemnent of ^{14}N with ^{15}N and ^{12}C with ^{13}C. The ^{15}N and ^{13}C substitutions aid in spectral simplification by expanding NMR resonances along the much wider (ca. 10-20 times larger) spectral widths. Multidimensional NMR techniques have been used for example for the study of the structure of human Interleukin-4 (Powers et al., 1992), and for many other systems (Fesik et al., 1990).

Areas of intense investigation and biological importance are studies of protein-protein interactions, protein folding and receptor-ligand interactions. In order to understand protein folding much work has been done on isolating and studying protein folding intermediates. Hydrogen exchange experiments, with the aid of NMR, have been done to monitor the exchange of hydrogens between a protein and its solvent. Exchange experiments help to understand how protein folding protects groups from solvent contacts. Exchange labelling has been used for example to understand allosteric structure changes in hemoglobin, and folding intermediates in pancreatic trypsin inhibitor (Englander et al., 1992).

Structural studies of intact complexes that bind tightly $K_d > (10^5-10^6)$ can be carried out by direct NMR methods, if the total molecular weight is not too high. Under normal conditions complexes of antibody protein with peptide antigen are too large to study by 3D or 4D NMR techniques. Limited structural information has been obtained on systems with molecular weights of up to 50 kD, in the case of a peptide antigen bound to a F_{ab}' fragment of a monoclonal antipeptide antibody (Tsang et al., 1992)(Cheetham, 1989).

For receptor molecular weights above 25 kd it is possible to study the conformation of weakly bound ligands such as peptides using Transfered NOESY NMR (Tr-NOESY). If

a ligand is in fast exchange with a large receptor system, the structure of the bound ligand can be monitored by NMR. For example this technique has been used to study the bound structure of Troponin I (peptide) to Troponin C (Campbell et al., 1991); C-terminal peptide of hiruden bound to thrombin (Ni et al., 1990); a dodecapeptide sequence of nicotinic acetylcholine receptor (nAchR) to an α-Bungarotoxin (Basus et al., 1993); and a G protein peptide bound to rhodopsin in the dark and light (Dratz et al., 1993).

2D and higher dimensional NMR has been a great aid in the study of fragments of larger proteins. For larger proteins, solution NMR falters in that these bigger biomolecules do not reorient rapidly enough in solution to sufficiently narrow resonances. Solid state NMR can be an effective tool for providing structural information on slowly reorienting biomolecules. Solid state NMR has no advantage over solution NMR for simplification of the spectral complexity of a system, however solid state NMR is able to obtain information on broad anisotropic resonances of large biomolecule systems not accessible to liquids NMR. These broad anisotropic resonances contain dipolar and chemical shift anisotropy interactions which depend on the structure and orientation of the system which are normally averaged out in solution. Solid state NMR has been used for example to study structural features of gramicidins (Nicholson et al., 1991); rhodopsins (Smith et al., 1988);

and phage coat proteins (Opella et al., 1987).

The first project to be described in this thesis explores the use of solid state NMR and aims to develop methods to recognize metal ion binding sites in proteins. It is often difficult to carry out spectral assignments in solid state NMR. One of the most general assignment techniques is specific isotopic substitution. One can carry out isotopic substitution biosynthetically, however, most proteins have many different amino acids of each type. Shift reagents are one possible way to discriminate between solvent accessible and inaccessible residues and thereby gain information that may aid in resonance assignments. The following chapter investigates the possible use of paramagnetic shift reagents in solid state NMR.

CHAPTER 2

PARAMAGNETIC SHIFT REAGENTS IN SOLID STATE NMR

Paramagnetic shift reagents have been used extensively in liquids NMR for spectral simplification, resolution enhancement, conformation determination, and assessment of enantiomeric purity (Sievers, 1973). Paramagnetic reagents have also been used in solids NMR but to a much more limited degree. The primary application of paramagnetic compounds in solids NMR has been to lower T1 relaxation times (Wenzel, 1987). Paramagnetic species are also potentially useful as shift reagents in the solid state, although spectral broadening effects of paramagnetic reagents can be quite large in the solid state. Well resolved spectra of several nuclei have been observed in solid state NMR in the presence of paramagnetic species. For example solid state carbon NMR spectra has been reported for praseodymium and europium acetates (Chacko et al., 1983). The static solid state ^{13}C NMR of praseodymium and europium acetates show broad resonances but magic angle spinning (MAS) partially averages the inhomogenous line broadening, narrows the linewidths of the resonances and improves the resolution. Another example of paramagnetic effects in solid state NMR spectra is the comparison of ^{119}Sn spectra of diamagnetic $Ln_2Sn_2O_7$ and its paramagnetic counterpart $Sm_2Sn_2O_7$. The average chemical shift of the ^{119}Sn resonance of $Sm_2Sn_2O_7$ is shifted downfield

compared to $Ln_2Sn_2O_7$ and this is believed to be due to the influence of the unpaired electrons of samarium (Cheetham, et al. 1987).

Paramagnetic ions have been useful as probes of metal ion binding sites in biology. There has been much work done to understand the coordination geometry of metal ion binding sites, and structural changes caused by metal ion binding to proteins. In particular calcium binding has a very large number of regulatory functions and carboxyl groups are often found to coordinate calcium in biological systems. Calcium binding proteins such as calmodulin and parvalbumin have a helix-loop-helix moiety called the EF-hand which binds calcium for regulatory purposes (Strynadka et al., 1989). However, there are very few spectroscopic probes of calcium binding sites. Certain paramagnetic ions have been used as spectroscopic probes to substitute for calcium ions and are often found to mimic calcium sufficiently to impart biological function (Ariki et al., 1986).

Since there are so few spectroscopic probes of calcium binding, techniques such as monitoring volume changes upon calcium uptake have been utilized. Studies with tetracarboxylate ligands showed that the size of the volume change can be related to the number of oxygen donor atoms in a metal ion complex, and to differences in structure of the liganding molecule (Kupke et al., 1989). It was found that volume changes can be used to monitor changes in the

coordination geometry of the metal ion binding site. It was seen in peptide analogues of binding loop II in the bovine calmodulin EF-hand that volume differences occurred in the coordination of lanthanide ions even with a minimal interchange of a carboxyl or amide group in the sequence (Kupke et al., 1989). Volume changes while somewhat informative cannot provide structural detail. In the vast majority of biological macromolecules of interest suitable crystals are not available to allow structure determination by x-ray diffraction. In addition, most biological macromolecules are not amenable to study by liquids NMR methods due to their large size, slow tumbling rate, low solubility or membrane binding. Solid state NMR has been a useful tool to provide structural information on immobile proteins including those in membranes (Smith et al., 1992). Chemical shift tensors from chemical shift anisotropy may provide orientational information on protein functional groups in membranes and have been used to study bacteriorhodopsin and bacterial phage coat proteins (Smith et al., 1992). For example solid state NMR of biosynthetically labelled ^{13}C aspartic acid sidechains in BR show resolved resonances, but assignment of the aspartic residues is difficult. Site specific mutagenesis of aspartic acids to asparagine or other residues can be used to shift or remove specific resonances for assignment purposes. However modifications of key residues in the

also alter the protein structure and indirectly alter the chemical shift of many residues due to electronic effects of nearby groups, changes in the tertiary structure, local pH effects, and local charges in the protein (Engelhard et al.,1989). Shift reagents that bind to carboxylic acid groups may help in the identification of specific surfaces accessible aspartic or glutamic residues in proteins by shifting effected resonances and simplifying the solid state NMR spectra.

Solid State NMR Theory

It is useful to start the treatment of NMR theory with the solid state since all the possible interactions are present, some of which are averaged to zero by rapid motions in liquid state NMR. The main interactions of magnetic nuclei are Zeeman, direct dipole-dipole, magnetic shielding, spin-spin coupling, and quadrupolar.

A Hamiltonian describing the sum of these interactions is:

$$H = H_z + H_D + H_{CS} + H_{SC} + H_Q \qquad 1$$

The Zeeman interaction (H_z) occurs between the magnetic moment of the nucleus and the large applied magnetic field (B_0). The interaction occurs with nuclei that have odd atomic number or odd mass. The Hamiltonian of the Zeeman interaction is:

$$H_z = -\gamma \hbar H_0 |_z = -g_N \beta_N B_0 |_z \qquad 2$$

where:

γ = magnetogyric ratio

g_N = nuclear g factor

β_N = Bohrmagneton for nucleus of interest

$|_z$ = is the nuclear spin quantum number

B_0 = is the magnetic field

The dipolar interaction (H_D) via from dipole-dipole interactions between nuclei. The dipolar interaction between an isolated spin pair I and S can be described as the vector product of the spin vectors $\vec{I}$, and $\vec{S}$. The strength of the interaction is scaled by the magnetogyric ratios (γ_1, γ_s):

$$H_D = H_{IS} = \gamma_I \gamma_S \frac{\hbar^2}{r_{IS}^3} \vec{I} * \hat{D} * \vec{S} \qquad 3$$

where D is the dipolar coupling tensor, (to be explained later) and r_{IS} is the internuclear distance.

Dipolar interactions tend to be most important for nuclei with a spin equal to half that have no quadrupole moment. Dipole-dipole interactions have an orientation dependence of $(1-\cos^2\theta)$ in the magnetic field as will be discussed later. The angle, θ, is between the magnetic field and the vector connecting the two interacting dipoles. The dipolar interaction is independent of the external field

(B_0) and falls with ($1/r^3_{IS}$) distance dependence.

The chemical shift interaction (CS) is due to the shielding of the nucleus by surrounding electrons. The Hamiltonian is:

$$H_{CS} = \gamma_I \hbar I \hat{\sigma} \vec{H} \qquad 4$$

where $\hat{\sigma}$ is the chemical shift tensor (which is explained later) and I is a spin.

The CS interaction, which is called chemical shift anisotropy (CSA) in solids, is linear in the applied field. Since CSA is a measure of the shielding effects of the surrounding electrons, CSA can be highly sensitive to local geometry. This makes the CS interaction a valuable tool in the solid state.

The spin-spin coupling interaction between two spins I and S is represented by:

$$H_{SC} = \vec{I} * \hat{J} * \vec{S} \qquad 5$$

where $\hat{J}$ is the spin coupling tensor (which is explained later). The spin-spin coupling is a field independent interaction and is normally of smaller magnitude than other magnetic interactions in the solid state.

The quadrupolar interaction occurs only for spins with over a half spin quantum number. This interaction occurs between the quadrupole moment eQ and the electric field

gradient around a nucleus and can be described as:

$$H_Q = \vec{I}\hat{Q}\vec{I} \qquad 6$$

where Q is the quadrupolar tensor of spin I. If H_Q is present it tends to be a major interaction which dominates NMR relaxation and the appearance of the NMR spectrum. Overall the dipolar, CSA, and quadrupolar interactions typically play the most important roles in solid state.

All the interactions discussed, D, σ, J, and Q can be described by tensors (3x3 matrices). The (3x3) matrices which describe the tensors can be diagonalized by rotation of the reference frame to give 3 principal diagonal elements which simplify the description of the interactions. For example for the chemical shift tensor:

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \rightarrow \begin{bmatrix} \sigma_{xx} & & \\ & \sigma_{yy} & \\ & & \sigma_{zz} \end{bmatrix} \qquad 7$$

Interactions which are isotropic in solution are often anisotropic in the solid state. Anisotropic interactions in the solid state depend on orientation. The principal tensors σ_{11}, σ_{22}, σ_{33} describe the shielding of the nucleus along different molecular axes and can be used to study structural aspects of nuclei in the solid state. For example, it is known that the NMR chemical shift of carbon in a carbonyl group is very sensitive to its orientation in

the magnetic field. If the (C=O) direction of the carbonyl group is perpendicular or parallel to the magnetic field, the chemical shift may differ by 200 ppm (Fyfe, 1983).

NMR studies of single crystals which have very limited ranges of molecular orientations must be carried out to deduce the molecular orientations of the tensor elements. Typical single crystals have sharp orientation-dependent NMR lines. In powdered samples all possible molecular orientations are present and a broad NMR powder pattern is observed over the full range of the CSA limits for the noncrystalline solid. A broad powder pattern makes measurement and use of the CSA more difficult (see below).

Magic angle spinning (MAS) NMR is quite frequently used in the solid state. The magic angle (54.44°) is the angle between the spinning axis of the sample and the magnetic field where. At this special angle the dipole-dipole interaction term $1/2(3\cos^2\text{-}\theta)$ goes to zero. Dipolar interactions are averaged to zero if the sample is spun at the magic angle at twice the magnitude of its CSA. By spinning the sample at a frequency less than that of the maximum range of the chemical shift anisotropy, the broad powder pattern of a resonance is broken into a central resonance and sidebands seperated by the spinning frequency. The spinning sideband intensity profile and the value of the spinning frequency can be used to determine the values of the principal chemical shift tensor elements (Herzfeld, and

1980).

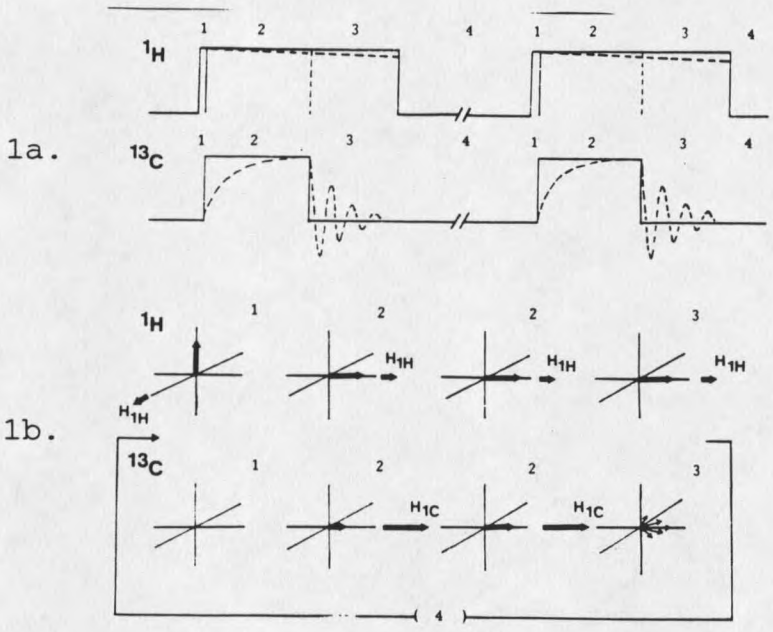
Cross polarization-magic angle spinning (CPMAS) is used in solid state NMR, to observe dilute nuclei such as ^{13}C. Spinning at the magic angle is used to increase resolution and cross polarization is used to transfer magnetization from abundant higher sensitivity nuclei (typically protons) to the dilute lower sensitivity nuclei, such as ^{13}C. Cross polarization increases sensitivity by the ratio of the magnetogyric moments of the abundant and rare spins (γ_H/γ_X).

The pulse sequence of ^{13}C CPMAS is shown in figure 1. In step 1 ^{1}H and ^{13}C magnetizations are rotated to the y-axis. Both magnetizations are spin locked along y during step 2. By adjusting the magnitudes of the spinlock fields of H_{1H} and C_{13C}, the Hartmann-Hahn condition is satisfied where the ^{1}H and ^{13}C magnetizations precess at the same frequency about the spin lock field. Under the Hartman-Hahn condition the proton magnetization is transferred to the ^{13}C magnetization to enhance sensitivity. In the Hartman-Hahn condition the ^{1}H and ^{13}C spin lock fields satisfy the relation:

$$\gamma_{1_H} H_{1_H} = \gamma_{13_C} H_{13_C} \qquad 8$$

Since the gyromagnetic ratio of protons and carbons is $\gamma_H/\gamma_C=4$, full cross polarization produces a fourfold gain in sensitivity for ^{13}C detection.

Figure 1. a. Pulse sequence used for cross polarization of ^{1}H to ^{13}C. b. Magnetization behavior of the ^{13}C and ^{1}H spins resulting from the pulse sequence in a (Fyfe, 1983).

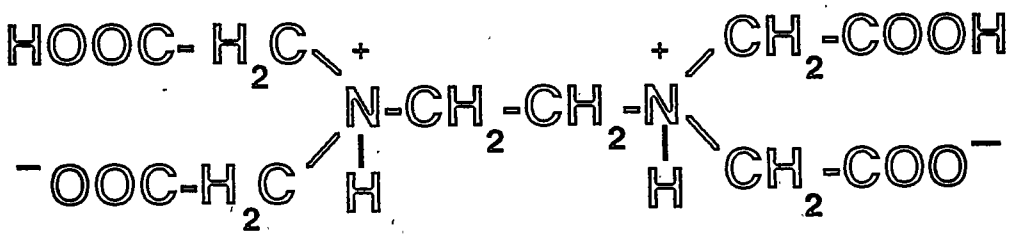


If sufficient spinlock time is allowed the ^{13}C magnetization has built up by step 3 and the carbon FID is recorded. The ^{1}H field is usually kept on for proton heteronuclear decoupling to remove C-H coupling from the ^{13}C spectrum and therefore greatly simplify the spectra. Step 4 is the relaxation delay before the next sequence of pulses. In addition to the magnetization enhancements the experiment also benefits from the recycle time depending on the relaxation of the proton spin system which normally is much faster than ^{13}C spins. This dependence on proton T_1's results in great savings in experimental time and enhanced sensitivity.

EDTA as a model paramagnetic metal chelator

Because of the interest and biological significance of metal ion binding in membranes and large proteins, we have sought to investigate metal binding to model compounds in solid state NMR using paramagnetic shift reagents. We have used EDTA as a model system as the first steps toward study of the solid state NMR behavior of paramagnetic metal ions complexing to carboxyl groups.

Ethylenediamine tetraacetic acid (EDTA) is a chelating agent which acts as a multidentate ligand to coordinate metal ions.



EDTA

The coordination number of EDTA in solution is typically either five or six ligand groups per metal ion (Nuttall et al., 1977). Coordination numbers up to ten have been found for EDTA-metal complexes in the solid state by X-

ray crystallography (Cotrait, 1972). Four carboxylate groups and two nitrogens are available in the EDTA molecule to coordinate metal ions. The lone electron pairs of carboxylate oxygens and the imide nitrogens can form coordinate bonds to the metal ions. The final coordination number and bonding strength depends on the radius and the electronic charge of the metal ions. Larger radius metal ions such as La^{+3} and Zr^{+3} favor a higher coordination number since EDTA cannot fully encircle the larger radius metal ions. With larger metal ions H_2O molecules usually bind to the metal ion in EDTA complexes at additional coordinate sites. Examples of these variations can be seen by comparing the structures of La^{+3}, Fe^{+2}, and Mg^{+2} EDTA chelates. La^{+3} forms a ten coordinate complex which is six coordinate with EDTA and four coordinate with H_2O. Fe^{+2} and Mg^{+2} form six and seven coordinate complexes with EDTA respectively, with one H_2O molecule coordinated to each metal ion (McConnell et al., 1978).

In the present work we used ^{13}C CP/MAS solid state NMR to determine the average chemical shift and to estimate the chemical shift anisotropy for natural abundance ^{13}C carboxylate groups of several metal:EDTA chelates. Molecular tumbling is fast in solution and tends to average out magnetic anisotropy information. However, in the solid state magnetic anisotropy along different directions in space gives rise to different chemical shifts. Aime et al

previously reported a solid state CP/MAS NMR study (Aime et al., 1986) to investigate the ^{13}C solid state NMR of EDTA complexes with the diamagnetic ions, Ca^{+2}, Mg^{+2}, and La^{+3}. Their results show that different diamagnetic metal ions shift the EDTA carboxylate resonances to different extents. In the present work we compared the effects of the diamagnetic ions Ca^{+2} and La^{+3} to the paramagnetic ions Sm^{+3}, Eu^{+3}, and Pr^{+3}, on the solid state ^{13}C CPMAS NMR spectra of the carboxylate groups of EDTA.

Our studies show that Eu^{+3} and Pr^{+3} broaden carbon resonances in the solids excessively for most purposes. However, Sm^{+3} can be used to detectably shift the ^{13}C resonances of metal bonded carboxylate ligand groups with modest line broadening. Thus, samarium ions are expected to be useful in solid state NMR to distinguish carboxylate groups in complex macromolecules that are accessible to cation binding from those that are inaccessible.

Experimental and Methods

Metal:EDTA chelates were prepared by adding stoichiometric (1:1 molar) amounts of metal chlorides (Gold Label grade from Aldrich or Strem Chemical Co.) to aqueous solutions of Na_2EDTA(Sigma Chemical Co.) Calcium chloride was added at ratios (mol metal/mol EDTA) of (2:1), 1:1 and 0.5 metal/EDTA. Sm:EDTA was prepared in a 0.5 and 1:1 metal/EDTA ratio. NaOH was added to adjust the pH to 6.5-7.0. The disodium salt of EDTA was used to prepare the

metal chelates because it had higher water solubility than H_4EDTA. The resulting solutions were dried under a vacuum at room temperature, followed by several hours drying in an oven at 80°C.

NMR experiments were performed on either a 200 MHz Chemagnetics Solid State CMC200 NMR spectrometer or a Bruker AM-500 MHz spectrometer operating at ^{13}C frequencies of 50 MHz and 125.7 MHz respectively. Samples were recrystallized from warm water, ground finely in a mortar, packed in zirconia rotors, and run at room temperature with spinning speeds of about 3.5 kHz or 5.0 kHz, respectively in the two machines. A standard CP/MAS pulse sequence with a contact time of 1 ms and a pulse delay of 20 sec was typically used. Four thousand free induction decays were typically collected per experiment and Fourier transformed with an exponential line broadening of 10-40 Hz. The aromatic ^{13}C resonance of hexamethylbenzene (HMB) was used as an external reference at 132.1 ppm. An interrupted decoupling pulse sequence was used in some experiments to suppress the methylene region in order to more easily observe higher spinning sidebands from the carboxylate region (Fyfe, 1983).

Spectral fitting of the CSA spinning side band intensities was accomplished using the program Spectracalc (Galactic Industries Corp.). Low order 3 or 5 point Savitsky-Golay smoothing functions were first applied to the NMR data. Each spinning side band was then expanded on the

x-scale, and a sixth order polynomial baseline correction was applied. A lineshape filling routine was applied to each partially resolved side band doublet. Initial guesses were applied manually by estimating the peak heights and widths, and the percentage of Gaussian or Lorentzian character of the peaks. Peak areas for each of the side bands were normalized by the area of the centerband.

Fourier transform infrared (FTIR) spectroscopy was performed with a Nicolet DX/FTIR. KBr pellet samples were prepared and the spectra were averaged for 10 scans at room temperature.

Crystallization was accomplished by dissolving 600-1000 mgs of Sm:EDTA in 20-30 mls of warm H_2O, The samples were placed in a lab oven at 50° C and the temperature slowly decremented over a one day period. Large crystals were produced by this method. The crystals were washed gently with water several times before mounting. X-ray crystallography was performed on a Nicolet (Siemens) model R3m/E automated diffractometer. A Patterson synthesis was used for the Sm^{+3} position and the structure was refined using the SHELXTL program package (Sheldrick, 1983).

Results

Many of the EDTA complexes gave significantly sharper CPMAS NMR spectra after the samples were recrystallized before being ground into microcrystals (data not shown). The beneficial effect of recrystallization was especially

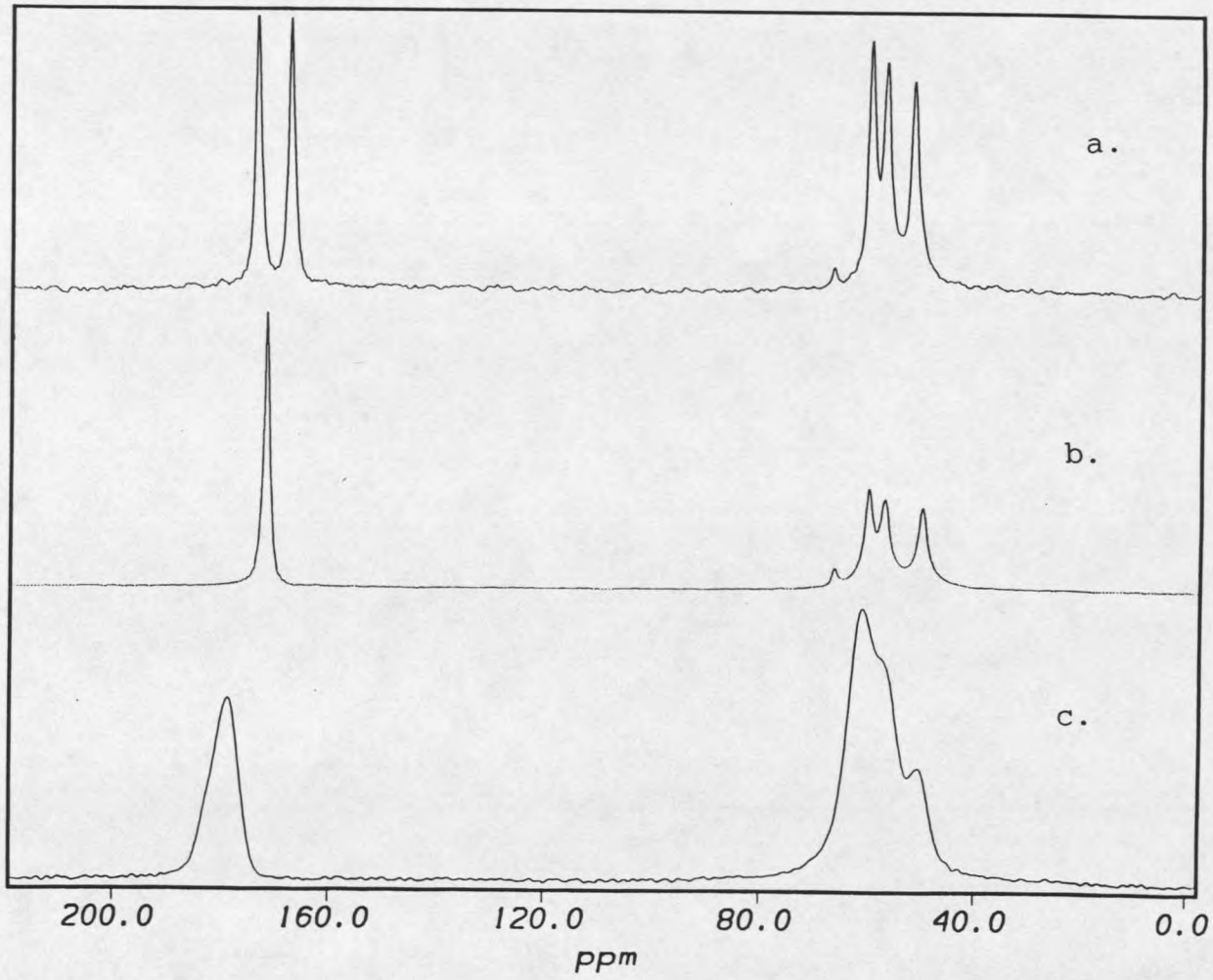
marked for the rare earth complexes and may have been due to the presence of a range of hydration states before recrystallization and a selection of hydration states by recrystallization. Natural abundance CP/MAS ^{13}C NMR spectra of microcrystalline H_4EDTA, Na_2EDTA, and Ca:EDTA are shown in figure 1. The methylene regions show similar resonances between 50 and 75 ppm with four overlapping peaks and little information could be drawn from this part of the spectrum. The carboxyl region of H_4EDTA shows two peaks of equal intensity at 174.7 and 168.8 ppm (figure 2 and table 1).

Table 1.) ^{13}C Chemical shifts of carboxyl groups in microcrystalline metal:EDTA chelates.

Compound	Resonance	Average
H_4EDTA	168.7 174.7	
Na_2EDTA	172.7	
Ca:EDTA	180.5	
La:EDTA	179.9 181.3	180.6
Sm:EDTA	180.0 185.5	182.8

The two carboxyl peaks in H_4EDTA are due to two of the carboxyl groups being protonated and two of the carboxyl groups being unprotonated (with two protons on nitrogen). The spectrum of Na_2EDTA shows one carboxyl peak at 172.7 ppm, and all four carboxyl groups are unprotonated. The Ca:EDTA chelate shows a broader resonance shifted downfield relative to Na_2EDTA at 180.5 ppm. The methylene and the carboxylate chemical shifts agree with

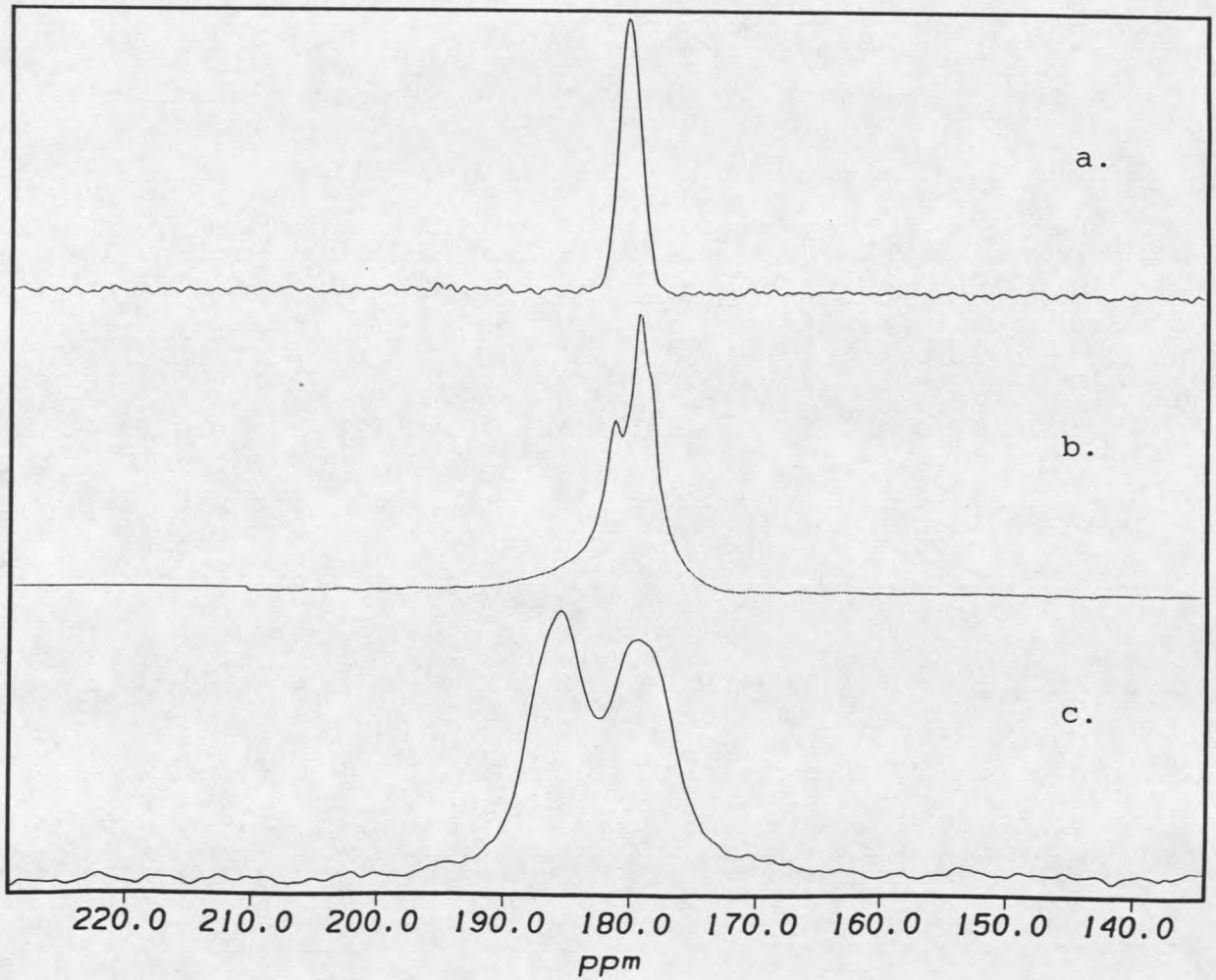
Figure 2. ^{13}C CP/MAS solid state NMR of a.) H_4EDTA b.) Na_2EDTA c.) 2:1 Ca:EDTA



values reported where available (Aime et al., 1986). Solid state NMR spectra of the carboxyl region of the Ca^{+2}, La^{+3}, and Sm^{+3} EDTA chelates are shown in figure 3 and the carboxyl chemical shifts are listed in table 1. La:EDTA shows fine structure in the carboxyl resonance compared to the Ca:EDTA. Note that the Ca:EDTA was not recrystallized and this may have led to a broader spectrum.

The Sm:EDTA chelate shows two carboxyl resonances separated by about 5 ppm, the resonances are broader than the diamagnetic complexes and are shifted 9 and 14 ppm downfield relative to Na_2:EDTA. The spectra for the Eu^{+3} and Pr^{+3} EDTA chelates show very broad carboxyl resonances (data not shown) as previously reported for the acetate complexes (Ganapathy et al., 1986). As in the Eu^{+3} and Pr^{+3} acetate cases there are large paramagnetic shifts of the carboxyl resonances in the Pr^{+3} and Eu^{+3} EDTA complexes, but the broadness of the peaks leads to very low signal intensities. Therefore, Pr^{+3} and Eu^{+3} appear to have very limited applicability as shift reagents in ^{13}C solid state NMR. However, we found that Sm^{+3} gave a small but significant shift along with modest broadening and appeared to be potentially a useful paramagnetic shift ion in solid state NMR. The Sm^{+3} metal ion shifts the carbonyl resonance of EDTA modestly but detectably from 172.7 to 180 and 185.5 ppm with an average shift of 182.8, while the average La^{+3} shift is to 180.6 ppm.

Figure 3. ^{13}C CP/MAS spectra of a.) Ca:EDTA b.) La:EDTA c.) Sm:EDTA



The Sm^{+3} complexes show a much better signal to noise ratio than Eu^{+3} or Pr^{+3} complexes. In figure 4 the spectrum of a substoichiometric (0.5:1) Sm:EDTA complex shows a free EDTA resonance at 172.5 ppm and a (1:1) complex spectrum (see Figure 1,2). Similarly a Ca:EDTA (0.5:1) or (1:1) complex shows a free EDTA resonance at the same position of 172 ppm. This 172 ppm resonance corresponds to the carbonyl groups of unchelated starting compound Na_2EDTA. Importantly, this result indicates that Ca^{+2} and Sm^{+3} shift the carbonyl resonances of the directly coordinated groups only and do not induce intermolecular, nonspecific shifts in nonbonded carboxyls more distant from the metal in the solid.

The different appearance of the carboxyl species in figure 4 compared to figure 3 may be due to lack of recrystallization for the samples in figure 4 and a different hydration state for 0.5:1.0 complexes.

Previous solid state NMR studies indicate that carboxyl ^{13}C chemical shifts depend on the OCO bond angle (Jagannathan et al., 1989) (Veeman et al., 1984). The X-ray structure of Sm:EDTA does not appear to have been reported previously, although the structure for La:EDTA is known. Therefore we determined the Sm:EDTA crystal structure to assist in interpretation of the solid state NMR spectra. Figure 5 shows the x-ray structure of Sm:EDTA which reveals that Sm^{+3} is coordinated by four carboxylate oxygens, two nitrogens and three waters. The x-ray structure suggests

that the double carboxylate resonances in Sm:EDTA may be explained by the two sets of OCO bond angles in the samarium complex. The two sets of OCO bond angles are grouped near 124.3 and 125.7 degrees (table 2). From Jagannathan's work a ~ 15 ppm difference in the σ_{22} chemical shift tensor values would be predicted. In the Sm:EDTA case a σ_{22} value of ~ 11 ppm was observed and will be discussed later.

Table 2. OCO Bond angles of EDTA chelates in crystals.

Compound	$O_2C_4O_1$	$O_7C_{10}O_8$	$O_3C_6O_4$	$O_6C_8O_5$
Sm:EDTA	124.3	124.2	125.8	125.6
La:EDTA	124.7	124.6	124.2	125.1

The La:EDTA has a broad carboxylate peak at 179.9 ppm with a downfield shoulder at 181.3 ppm and 1-2 other shoulders (table 2). The broadening effect of La is thought to be partly due to a quadrupole interaction of ^{139}La (I=9/2, natural abundance of 99.9%) with ^{13}C. La:EDTA has a similar X-ray structure to Sm:EDTA. Both the Sm^{+3} and La^{+3} EDTA complex crystal structures show one metal per unit cell (Lind et al., 1965) (Koetzle et al., 1972) while the Ca^{+2} complex shows two (Bartlett et al., 1979). The 2-3 shoulders on the La:EDTA carboxyl resonance might be rationalized by the 3 slightly different OCO bond angles in the La complex (table 2) of 124.2, 124.7(2), and 125.1 as found in the x-ray structure.

Figure 4. ^{13}C CP/MAS NMR spectra of the carbonyl region of a.) 0.5 Ca:EDTA b.) 0.5 Sm:EDTA.

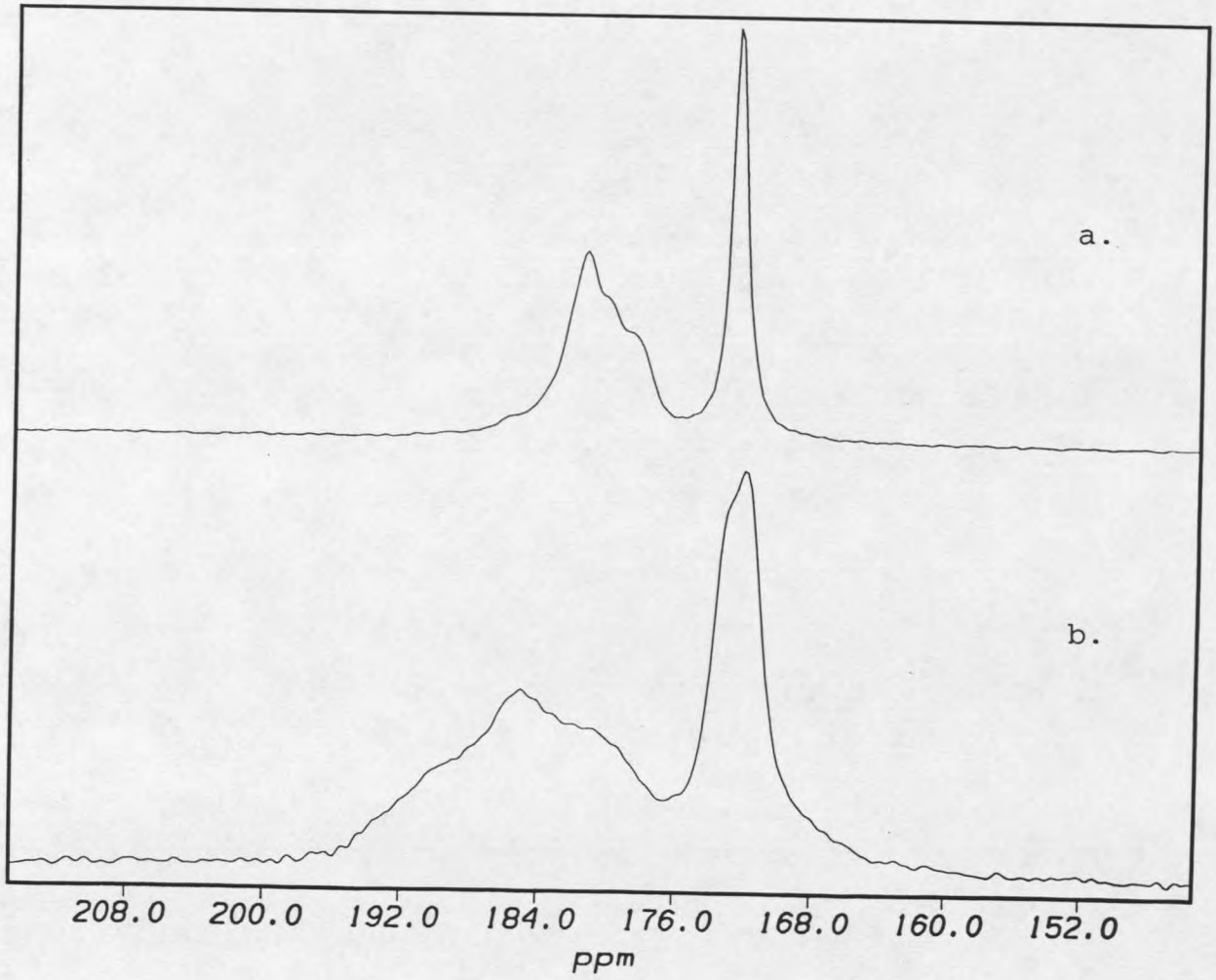
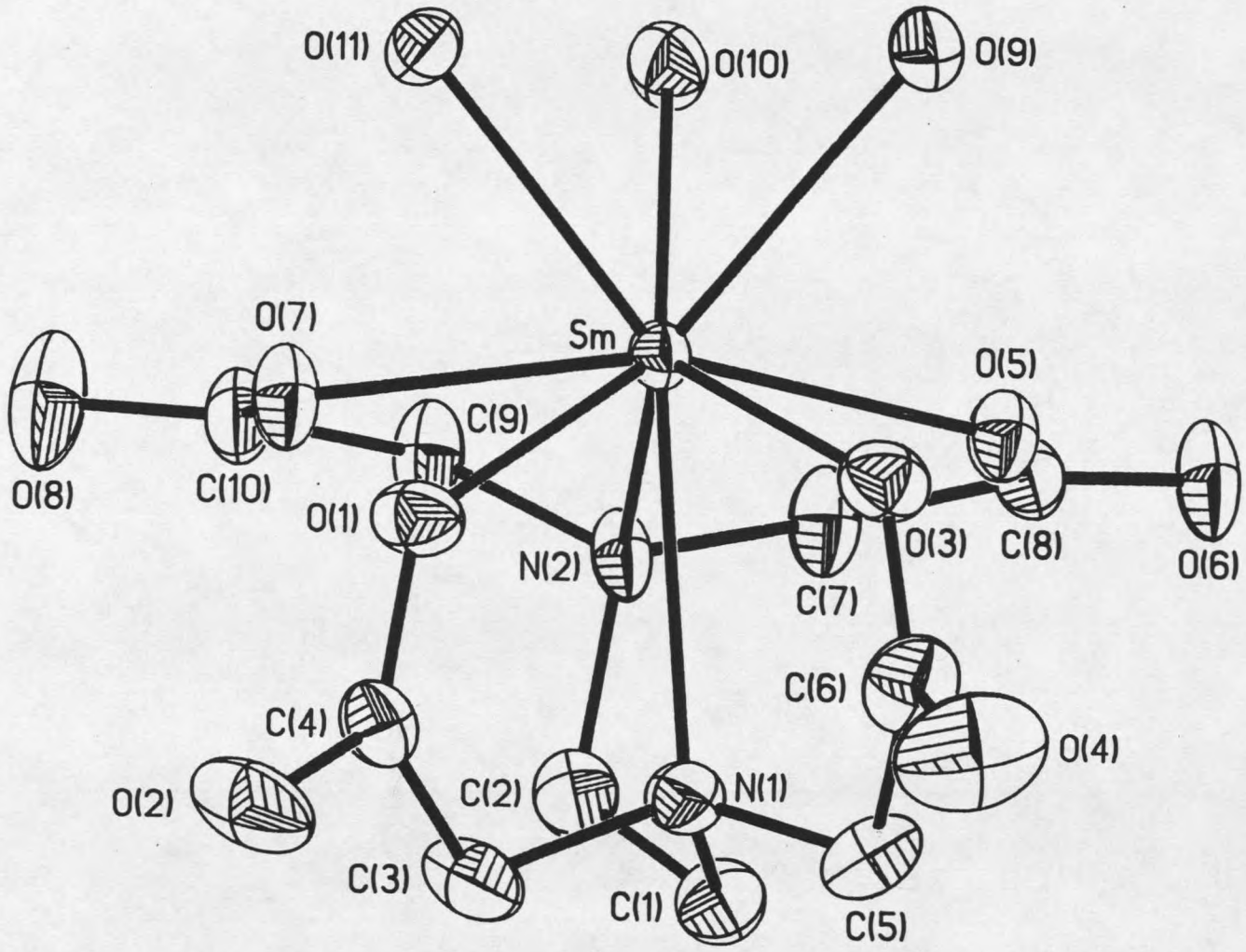


Figure 5. X-ray crystal structure of Sm:EDTA



Slower spinning in CPMAS NMR allows observation of an envelope of sideband intensities that depend on the size of the components of the CSA. The two types of sites in the Sm:EDTA appeared to have slightly different CSA as can be seen in the slightly different intensity profiles of the left and right CPMAS sideband spectra in figure 6. The pairs of peaks near the center band are best resolved and the different relative shifts of the higher order sidebands imply a somewhat different CSA pattern for the two types of sites.

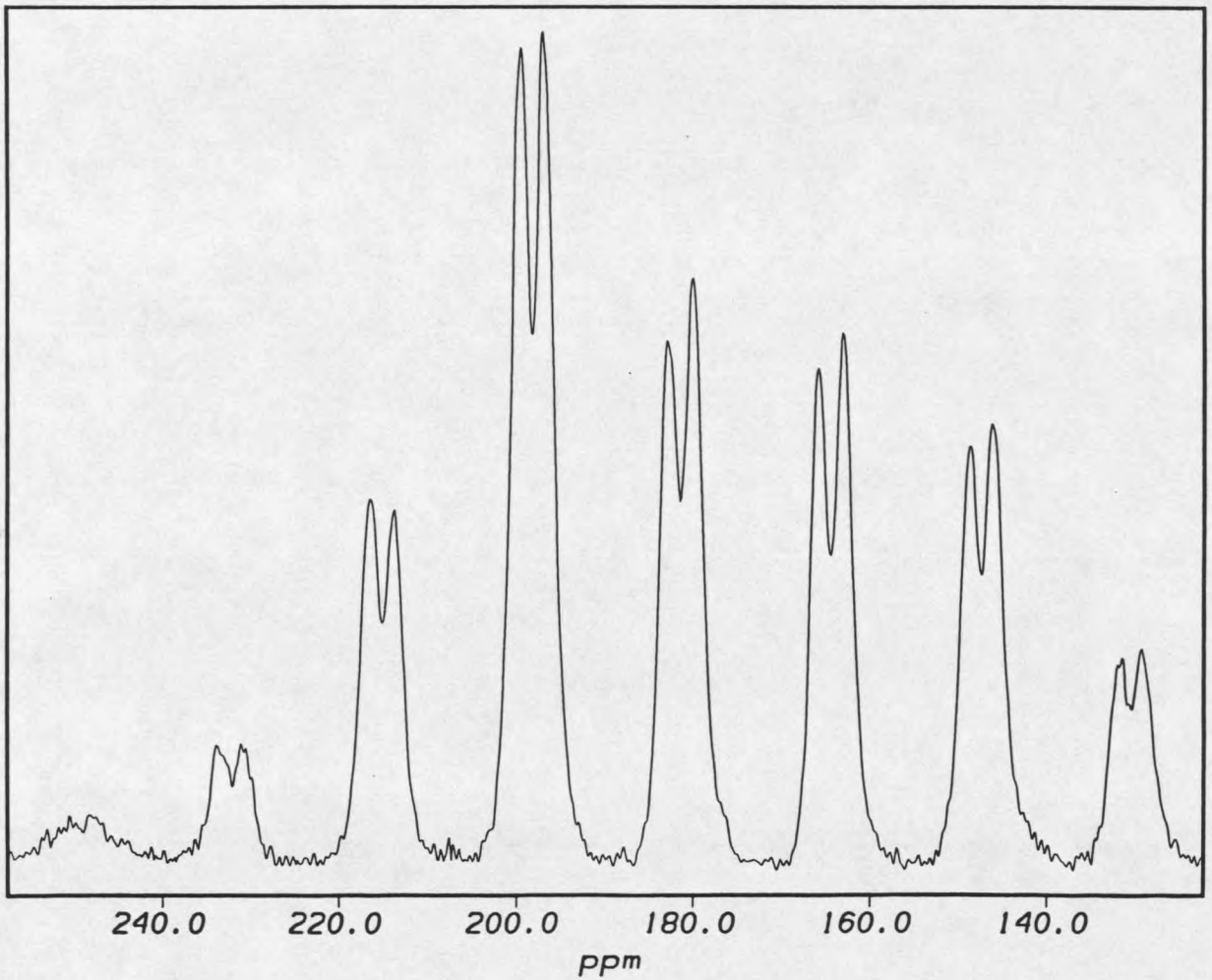
The CSA can be quantitatively described by a tensor and converted to an array of three principal elements on the diagonal of a 3x3 σ matrix in the shielding Hamiltonians (below).

$$H_s(i) = g_N^2 \beta_N^2 |_i \hat{\sigma}_i \vec{H}_0 \qquad 9$$

The above equation describes the orientation of the chemical shielding where σ_i is the second rank tensor. Principal elements of σ ($\sigma_{11}, \sigma_{22}, \sigma_{33}$) describe the shielding in different directions of space around a nucleus. The CSA is averaged in solution into isotropic resonances, however in the solid state due to diminished molecular motion the chemical shifts are anisotropic.

Different hydration states have been observed to give different chemical shifts of metal ion complexes in CPMAS (Ganapathy et al., 1986). We assumed that we had a single

Figure 6. ^{13}C CP/MAS NMR spectra of Sm:EDTA showing three orders of spinning sidebands at a spinning speed of 4200 Hz.



hydration state in the materials that were recrystallized. The x-ray crystal structures of recrystallized samples showed clear hydration states. If samples were not recrystallized the NMR spectra were broader which suggests that the resonances can be broadened by a mixture of hydration states.

The appearance of the spectra depends on the spinning speed ω_r and maximum anisotropy width $\Delta\sigma=\sigma_{33}-\sigma_{11}$. When $\omega_r >> \Delta\sigma$ the chemical shift anisotropy collapses to a single line located at the isotropic average $(\sigma_{11}+\sigma_{22}+\sigma_{33})/3$ isotropic chemical shift. In the other extreme case when $\omega_r < \Delta\sigma$ the spectra will contain additional lines spaced by the spinning frequency (spinning sidebands) that trace out the static CSA intensity profile within the σ_{11}, and σ_{33} CSA limits. The sidebands result from the refocussing of the magnetization after each rotor cycle. The intensity profile of the spinning sidebands can be used to obtain the magnitudes of the principal elements of anisotropic chemical shift tensor. The method of Berger and Herzfeld uses the spinning sideband intensity profile to extract $\sigma_{11}, \sigma_{22}, \sigma_{33}$ (Herzfeld and Berger, 1980).

If large single crystals are available and the crystal structure is favorable, CSA tensor elements can be connected to the molecular coordinate system in the crystal. There are many cases when sufficiently large crystals are not available, but magnitudes of the chemical shift tensor can

can be determined from the anisotropy of the NMR powder pattern measured on static samples. The disadvantage in using static powder patterns for determining the magnitude of the CSA is that the powder patterns may cover very large spectral widths and therefore have weak signals with low sensitivity. Furthermore, if NMR active isotopes are not enriched and samples are complicated chemically, the static powder pattern may be heavily overlapped.

We have used the method of Berger and Herzfeld to estimate CSA values from spinning sideband intensities in our solid state NMR studies of metal:EDTA complexes. We used hexamethylbenzene (HMB) as a test case since the CSA tensor values were well known (Mehring, 1983). In the case of HMB at room temperature, the two tensor elements σ_{22} and σ_{33} are averaged by the spinning of the plane of the phenyl ring, while σ_{11} coincides with the axis of molecular rotation, perpendicular to the plane of the phenyl ring in the solid. As seen in table 3 our results agree quite well with literature CSA values obtained on static samples of HMB (Mehring, 1983).

Table 4 shows the CSA tensor values of the carboxyl resonances for three sets of ^{13}C CP/MAS solid state spectra of Sm:EDTA calculated using the Herzfeld Berger method. In order to test the validity of the calculated CSA tensor values, with the calculated CSA tensor values (table 4), the

rotor spinning speed, and the isotropic chemical shift, were used in a MAS sideband

Table 3. CSA values of HMB aromatic carbons.

	σ_{11}	σ_{22}	σ_{33}
Experimental	-112	52	60
Literature	-112	56	56

Table 4. CSA values of sideband pattern of the left (LT) and right (RT) ^{13}C carbonyl resonaces of Sm:EDTA.

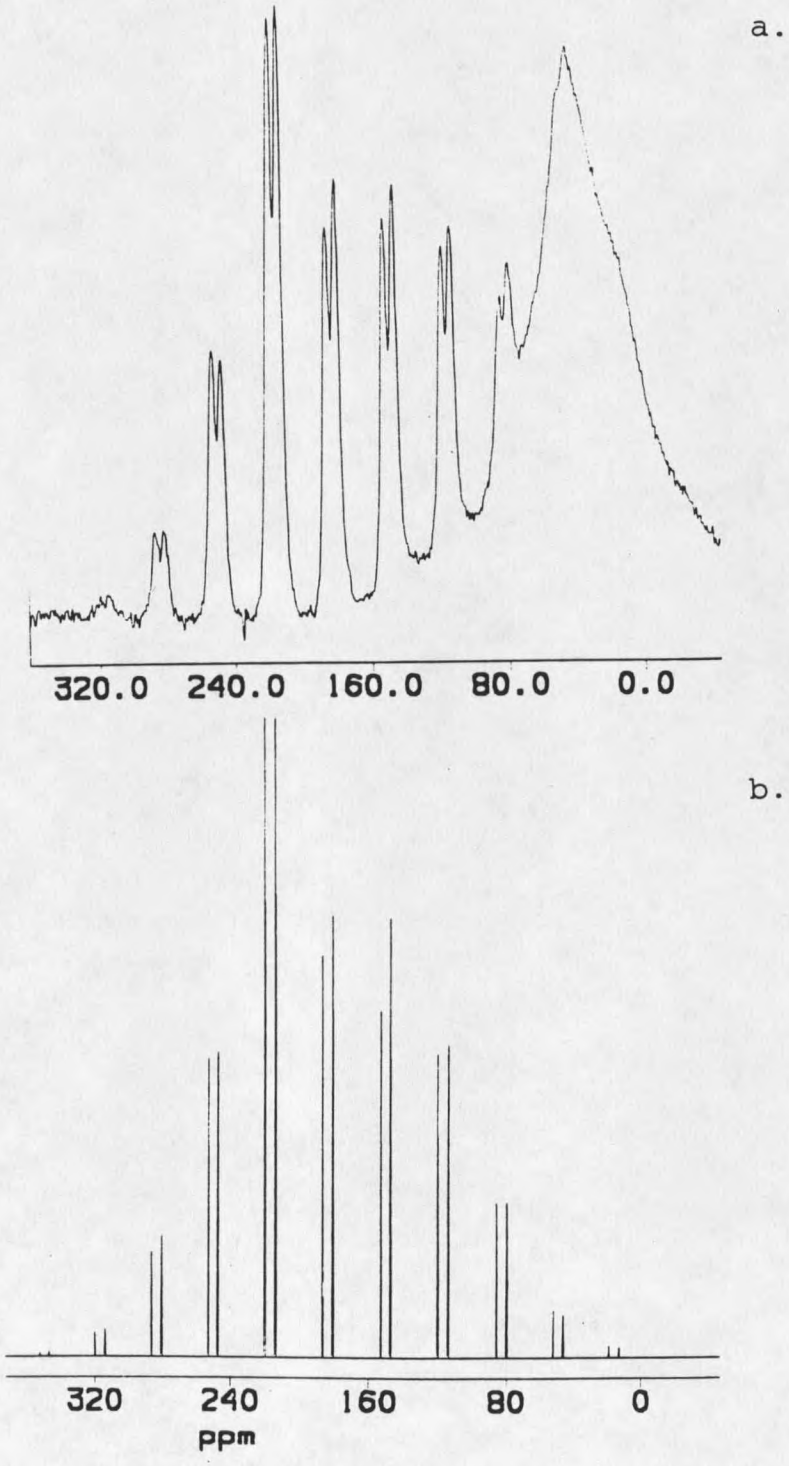
sample	resonance	σ_{11}	σ_{22}	σ_{33}
I	LT	-124	21	103
	RT	-119	13	106
II	LT	-108	16.7	92
	RT	-108	18.4	89
III	LT	-122.4	22.3	100.2
	RT	-106	11.5	95.2
AVG CSA	LT	-118.1	20	98.4
	RT	-111	14.3	95.7

simulation program in FTNMR (Hare Research). Figure 7a and 7b show the experimental and simulated MAS sideband

resonance pattern of Sm:EDTA respectively. As can be seen the simulated MAS sideband intensities agree quite well with the experimental values. From the comparison of the calculated Herzfeld Berger CSA tensor values and literature CSA tensor values for HMB, and the experimental and simulated MAS anisotropic resonance sideband intensities of Sm:EDTA, it is seen that the Herzfeld/Berger method provides quite reasonable CSA tensor values for these samples.

The σ_{22} tensor is known to be especially sensitive to the local structure in carboxylic systems. Jagannathan reported that the σ_{22} tensor can be related to the OCO bond angle for carboxyl carbons. His results showed that the σ_{22} tensor changed ~15 ppm +/-5 ppm for a one degree (out of 360°) change in the OCO bond angle in dicarboxylic acids. We tried therefore to relate the values of the tensor elements calculated from the spinning sideband intensities to the OCO bond angles of La:EDTA and Sm:EDTA complexes which are known from their x-ray structures. Jagannathan's results show an inverse relationship between σ_{22} and the carboxyl OCO bond angles. Smaller σ_{22} tensor values correspond to larger carboxyl OCO bond angles. In the case of Sm:EDTA the σ_{22} value of 14.3 would reasonably correspond to the carboxyl OCO bond angle set of 125.8° and 125.6° (table 2) while the σ_{22} tensor value of 20 would correspond to the OCO bond angle set of 124.2° and 124.3° (table 2).

Figure 7. a.) ^{13}C CP/MAS spectra of Sm:EDTA at a spinning speed of 4200 Hz. b.) Simulated MAS anisotropic resonance pattern from a. The large poorly resolved peak on the right id due to methylene resonances.



Upon repeated fits of CSA values on different data sets it was found that the difference in the average σ_{22} tensors for the two carboxyl carbon resonances in Sm:EDTA (table 4) was ~11 ppm. The 11 ppm difference in σ_{22} with in the +/- 5 ppm uncertainty would correspond to a one degree change in the OCO bond angle of the carboxyl carbon resonance.

The Herzfeld/Berger CSA calculation is quite sensitive to relative peak areas of spinning sidebands. It was seen that with the overlapping sideband doublets small changes in the baseline effected peak area determination sufficiently to effect the σ_{22} tensor value. Peak fitting of the main resonance of the La:EDTA complex proved to be more difficult due to the limited resolution of shoulders on the main resonance. Successful fitting of the more poorly resolved La:EDTA sidebands could not be reliably carried out.

Carbonyl vibration frequencies in the infrared can be monitored for changes in the C-O bond length upon the coordination of a metal ion. COO^- groups have an asymmetric stretch at 1550-1610 cm^{-1} and a symmetric stretch at 1400-1450 cm^{-1}.

Table 5. Infrared stretching bands of carboxylate groups in EDTA Chelates.

Compound	Asymmetric	Symmetric	(Asym-Sym)
EDTA	1702		
Na2EDTA	1629	1396	233
Ca:EDTA	1616	1396	220
La:EDTA	1594	1403	191
Sm:EDTA	1562	1410	152

Discussion

Proton decoupling was applied during the ^{13}C CP/MAS measurements which largely eliminates proton-proton dipolar interactions. There is no (^{13}C-^{13}C) homonuclear coupling due to the low natural 1.1% ^{13}C abundance. So these contributions are minimal. Fermi contact (scalar) and paramagnetic dipolar mechanisms play the major roles in paramagnetic shifting and broadening in ^{13}C-CPMAS solid state NMR. Diamagnetic and Fermi contact interactions that shift and broaden the resonances and which operate through the bond framework depend on the metal to ligand bonding. Any paramagnetic dipolar interaction between the electronic magnetic moment and the nuclear spin moment is a through space effect which depends on the position of the paramagnetic ion (Martin et al., 1980). It was important to show that only the carboxyl groups that were directly

bonded to the Sm^{+3} showed shifts in the NMR and that nearby carboxyl groups in the crystals that were not bonded to Sm^{+3} did not shift (figure 4).

The difference between the asymmetric and symmetric carboxylate frequencies in the IR decreases going down table 5 seem to indicate an increase in polarity of the oxygen to metal bond (Alcock et al., 1976). The change in the difference between the asymmetric and symmetric frequencies becomes smaller as the size and charge of the metal ion increases with Sm^{+3} having the smallest difference. Sm^{+3} has the biggest polarization effect on the oxygen to metal bond due to its size and charge effect which is reflected in the smallest difference between the asymmetric and symmetric frequencies in the IR.

We have seen that Sm^{+3} shifted the bound carboxylate carbon resonance an average of 2 ppm further downfield than the diamagnetic La^{+3}, and Ca^{+2} ions. This may be due in part to the large positive charge on Sm^{+3} which is reflected in the carboxyl frequencies in the IR (table 5) and in part to its paramagnetic property (Sievers, 1973). It was also seen that small chemical shift changes in the Sm^{+3} bound carboxyl carbon resonances were related to significant effects in the σ_{22} tensor element which is a sensitive probe of changes in the OCO bond angle in carboxyl systems. Both of these effects suggest that Sm^{+3} may be a useful probe to

differentiate accessible carboxyl resonances in solid state NMR.

It was seen that quite accurate relative peak areas are important to detect reliable CSA tensor information for different chemical environments. New techniques have been recently developed that may help in peak area determination and improvements in spectral resolution of paramagnetic solids in solid state NMR. De Groot has devised a peak fitting analysis procedure for MAS spectra which assumes simple lineshapes for the main and spinning side band resonances (De Groot et al., 1991). A new technique by Raleigh (Raleigh et al., 1992) uses double frequency decoupling which appears to be useful in cases where there is a wide spread of proton frequencies, common in paramagnetic solids. This technique was shown to increase the spectral resolution in C-13 MAS spectra in paramagnetic solids (Raleigh et al., 1992). Techniques such as these may help in the future use of CSA tensors and paramagnetic shifts as probes of structural and environmental changes due to metal ion binding in biological systems.

CHAPTER 3

NMR STRUCTURE DETERMINATION OF A MODIFIED EF-HAND PEPTIDE: A MODEL FOR A SUFACE LOOP OF BACTERIORHODOPSIN

Membrane proteins play central roles in a wide range of biological processes such as signal transduction, ATP synthesis and ion pumping. Knowledge of the atomic structure of biological macromolecules, when available, has been extremely useful to increase understanding of their mechanisms of action. However, structural information on membrane proteins has been very limited. Protein structures have largely been elucidated through the use of single crystal x-ray diffraction methods, however membrane proteins have been largely refractory to crystallization for x-ray studies.

Due to the difficulty of crystallization only three intact integral membrane proteins have been solved to or near atomic resolution by x-ray or electron diffraction and all three are from bacteria. The highly helical heterotrimeric structure of photosynthetic reaction centers (RCs) from purple bacteria has been determined to 2.3 Å by X-ray diffraction (Rees et al., 1989). The X-ray structure of bacterial porin from *R. capsulatus* has been solved with a

resolution of 1.8 Å revealing a transmembrane beta barrel (Weiss et al., 1991). Bacteriorhodopsin is a seven transmembrane helix protein that uses visible light to pump protons across its membrane to create an electrochemical gradient. Bacteriorhodopsin (BR) occurs naturally as 2D crystals in the plasma membrane of halobacteria and can be recrystallized into larger 2D crystals. The structure of BR has been revealed by electron cryo-microscopy on 2D crystals to a 3.5 Å resolution in a 2D projection onto the membrane plane and approximately 5.5 Å in 3D by image reconstruction from a series of tilted specimens (Henderson et al., 1990). The tilted images range over +/- 60° and a "missing cone" of diffraction information results in a loss of structural information and very poor definition of the BR surface loops on BR that connect the helices.

Solution NMR cannot provide detailed structural information on intact membrane proteins because of the low mobility in membranes and the large size of detergent complexes of solubilized membrane proteins (Clore et al., 1993). The very limited motional freedom of proteins in native or synthetic membranes leads to broad NMR resonances except for mobile tails (Smith et al., 1984). Solution NMR relies on rapid isotropic motion to collapse CSA and narrow NMR resonances. The dipolar and chemical shift anisotropy (CSA) interactions which are averaged out in solution NMR

are present in the solid state and can be used in studies of protein structure determination and dynamics (Cornell et al., 1980). The solid state NMR techniques may be helpful providing information of protein structure, but it is not currently practical to use solid state NMR as a sole source for 3D structure of any but the smallest membrane proteins (Smith et al., 1992).

While whole membrane proteins in their natural environment are not easily studied by x-ray diffraction and NMR, these methods have been successfully applied to structural studies of smaller domains of membrane proteins. For example, the transmembrane signal protein, aspartate receptor Che A, mediates chemotaxis behavior in E-coli. Che A is a dimer of two subunits and the structure of the aqueous aspartic acid binding domain outside the membrane has been solved by x-ray crystallography (Scott et al., 1993). The three dimensional structure of the extracellular portion of human class I (Bjorkman et al., 1987) class II major histocompatibility antigen HLA-DR1 has been determined by x-ray crystallography (Brown et al., 1993). The x-ray crystal structure of the complex between human growth hormone and the extracelluar domain of its receptor (hGHbp) was carried out at a resolution of 2.8 Å (DeVos et al., 1992). All of the above cases have been domains that normally contact the aqueous phase much like a water soluble

protein.

Multidimensional NMR has been used to determine the structure of the 79 residue membrane transversing subunit c from the F_0 part of ATP synthase (Moody et al., 1987). Two Dimensional NMR has been used in structural studies of hydrophobic transmembrane subdomains of Bacteriorhodopsin in organic solvents. Through the use of peptide synthesis, enzymatic cleavage of specific regions, and N-15 isotopic labelling, data from solution NMR in methanol and chloroform aided secondary structure determination of BR segments (Pervushin et al., 1991)(Barsukov et al., 1992). In selected cases a great deal can be learned about the structure of membrane proteins by using multidimensional NMR and/or X-ray diffraction on fragments of membrane proteins.

NMR and the Nuclear Overhauser Effect

Most of the structural information obtained to date from NMR of macromolecules comes from studies in solution. If motions in solution are fast ($1/\tau_c >> \sigma_{33}-\sigma_{11}$) they average away the CSA and give single lines at $\sigma_{ave}=(\sigma_{11}+\sigma_{22}+\sigma_{33})/3$. With sharp NMR lines much more information can be obtained from chemical shifts, NOEs, and spin-spin coupling. The NMR information which can be used in protein structure analysis in solution can be described by six parameters. Three of these; chemical shift (δ), spin-spin coupling constants (J), and areas of resonances aid in assigning resonances in NMR

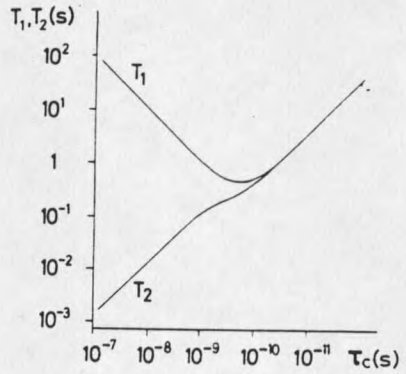
spectra. The other three parameters, spin lattice relaxation time (T_1), nuclear Overhauser effect (NOE), and spin-spin relaxation time (T_2), provide dynamic information. The NOE is also extremely useful to provide distance information and as such is the main source of information for NMR structure determination in biological systems.

The chemical shift is used in NMR to identify the position of the resonance of each nucleus. Chemical shifts are largely determined by local electron density variations due to electron density, nearby charges, dipoles, location of double and triple bonds, ring currents, and unpaired electrons. Predicting the chemical shifts of resonances and the shape and intensity of the resonance lines is difficult (Jardetzky, 1981).

The rate at which a rigid molecule tumbles in solution can be described by a correlation time (τ_c). This rate depends on such factors as the size of the molecule, viscosity and temperature. The relationship of T_1 and T_2 relaxation to correlation time for tumbling is seen for example in figure 8 (Wuthrich, 1986). Small molecules tumble at a fast rate ($\omega_0 << 1/\tau_c$), which is called the extreme motional narrowing region, where ω_0 is the Larmor frequency. This region is where T_1 and T_2 are equal and increase proportionally as the correlation time (τ_c) becomes smaller. In the case of large molecules and slow tumbling

$\omega >> 1/\tau_c$, which is called the spin diffusion region T_1 increases proportionally as τ_c becomes larger while T_2 decreases. In cases where $\omega_0 >> 1/\tau_c$, the small T_2 causes broadening of the resonances. The resulting broadening can cause difficulty in assigning resonances.

Figure 8. Logarithmic plot of T_1 and T_2 relaxation times versus the correlation time τ_c (Wuthrich, 1986).

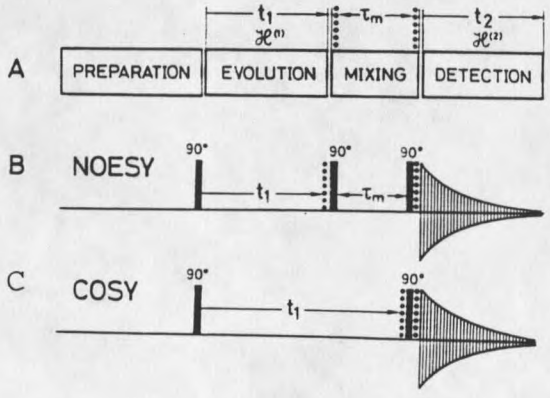


Line broadening can be increased by slower tumbling due to aggregation or increases in viscosity. As correlation time decreases for a small molecule, T_1 may increase to a point where long recycle times are needed for the recovery of the NMR signal intensity. However, it is the T_2 effect which tends to limit the size of macromolecules that can be effectively studied by NMR.

The relative complexity of the ^1H NMR resonances of a biomolecule can be observed in a simple 1D experiment. The usefulness of 1D NMR methods in protein structure analysis is quite limited because of spectral complexity due to many proton resonances in a small frequency window. This limitation of 1D NMR encouraged the development of 2D NMR. Two dimensional NMR is an extension of the 1D experiment by the addition of an extra variable delay (t_1) and one or more additional pulses. The end result is a NMR technique that spreads the ^1H resonances of a molecule over two frequency dimensions which helps in spectral resolution and assignments of resonances and increases the molecular size of biomolecules that can be studied.

The basic 2D NMR experiment is described in the following figure 9.

Figure 9. A is the general scheme for all 2D NMR experiments. B and C are two particular 2D NMR experiments. The dotted lines indicate the boundary of a mixing period or pulse (Wuthrich, 1986).



The preparation time is the delay to allow for return toward thermal equilibrium. This delay is followed by one or more rf pulses. The evolution period (t_1) is where chemical shifts are coded by variable xy magnetization amplitudes. The mixing time, if present, is where magnetization may be transferred among spins. During the detection period, FIDs are recorded in the time domain (t_2).

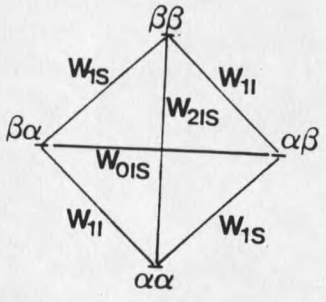
There are many variations of 2D NMR experiments and many of them can be tailored to observe specific connectivities. Two dimensional NMR experiments can be broken down into two general categories: scalar (spin-spin) and through space. The COSY and NOESY experiments typify the scalar and through space categories respectively. Both categories of 2D NMR experiments are used to assign the resonances to chemical groups in a molecule. Correlation spectroscopy (COSY) experiments are used for intraresidue connectivity, while the nuclear Overhauser effect spectroscopy (NOESY) experiments allow intraresidue and interresidue assignments.

Once spectral assignments have been accomplished, crosspeak intensities from 2D NOESY at appropriate mixing times can be related to internuclear distances. In order to understand the use of cross peak intensities it is useful to look at the two spin model which is the simplest system that exhibits NOE.

1D NOE Theory

Figure 10 shows the energy levels and transitions between levels for two spins I and S. Initially the 1D steady state NOE experiment will be considered where transitions of the spin S are saturated, and the I spin is measured. In order to exhibit an NOE effect the two spins must be close enough in space for dipolar relaxation to occur. The two spin transitions in figure 10 that give rise to NOE enhancements are the $\alpha\alpha$ to $\beta\beta$ (W_{2IS}) and the $\alpha\beta$ to $\beta\alpha$ (W_{0IS}) (Neuhaus and Williamson, 1989).

Figure 10. The energy level of a two spin system showing transition probabilities and spin states (Neuhaus and Williamson, 1989).

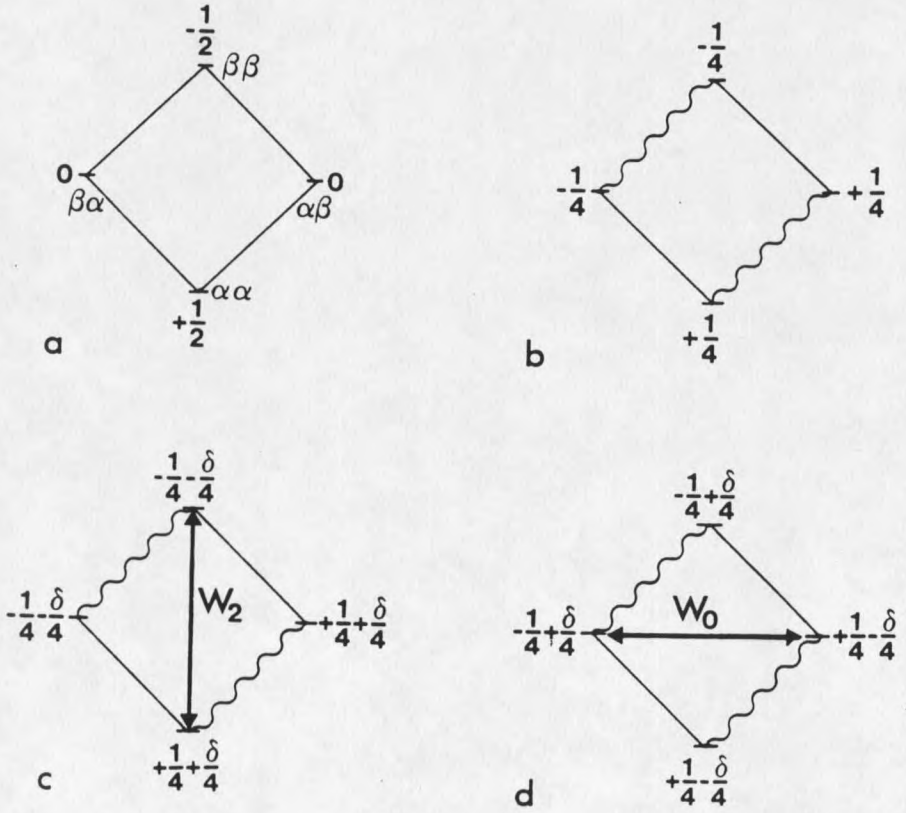


The W_{2IS} is known as the double quantum transition where both spins flip the same way and the W_{0IS} is the zero quantum transition where both spins flip the opposite way. The other W's, which are single quantum transitions, do not give

rise to NOE effects. Upon saturation of the single quantum transitions of the S spin, the populations of the irradiated levels are equalized and the resulting effects of this pertubation can be seen in figure 11. The populations of the levels are altered by the W_{2IS}, and W_{0IS} transitions which produce a positive and a negative NOE enhancement respectively. The W_{0IS}, zero quantum, transition corresponds to a frequency of $(\omega_I - \omega_S)$ which is the difference of chemical shift between the two spins. This difference would be in the range of 0 to 5 kHz range for ^1H at a magnetic field of 500 MHz. The $(\omega_I + \omega_S)$, double quantum, W_2 transition for ^1H in a 500 MHz (W_0) magnet would translate to a frequency of 1,000 MHz. The rates of the relaxation transitions W_0, W_1, and W_2 depend on the amplitude of the spectral density function at $W_1 - W_s \approx 0$, $W_I \approx W_S \approx W_0$, $W_I + W_S \approx 2W_0$ respectively.

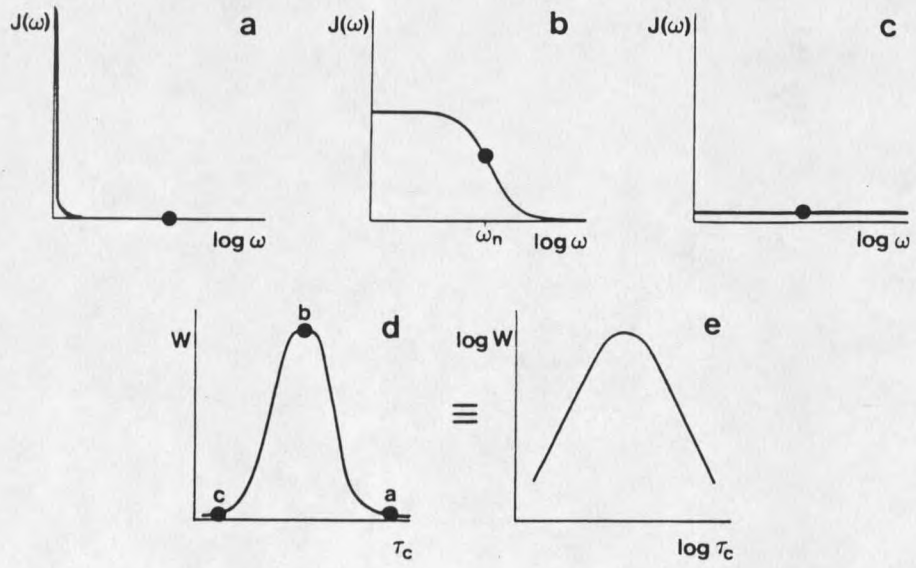
The spectral density function $J(\omega)$ describes the amplitude of magnetic fluctuation (power) available from a system's surroundings (lattice). If magnetic fluctuations occur at suitable frequency they help bring about the relaxation transitions (W_i) of interest. In figure 12 a, b, and c are the spectral density functions for slow, intermediate and fast tumbling systems in which the solid dot in each graph represents an observed frequency ω_n. The relationship between W and τ_c is more evident in d and e

Figure 11. The origin of the NOE in a two spin homonuclear system. Step a is the system at equilibrium where $\alpha\alpha$ is the most populated followed by $\alpha\beta$, $\beta\alpha$ and $\beta\beta$. Step b is the system after saturation of spin S where the populations of pairs of levels $N_{\alpha\alpha}$ $N_{\alpha\beta}$ and $N_{\beta\alpha}$ $N_{\beta\beta}$ are equalized. Step c shows the W_2 (double quantum) relaxation where ($\delta/2$) population units are transferred from the $\alpha\alpha$ to $\beta\beta$ state which results in a postive NOE enhancement of the I spin transition intensities. Step d shows the W_0 relaxation (zero quantum) where ($\delta/2$) population units are transferred from the $\beta\alpha$ to the $\alpha\beta$ state and thereby causing a negative NOE enhancement of the I spin transition intensities (from Neuhaus and Williamson, 1989).



(the log plot of d) where W may vary widely with respect to τ_c. For a small molecule which has a τ_c of > 10^{11} Hz the W_2 transition is much greater than W_0 and a positive NOE occurs. For slower moving molecules there is little or no spectral density at W_2, W_0 will dominate the relaxation and negative NOEs occur. At $\omega_0\tau_c \approx 1$ W_2 balances W_0 and zero NOE occurs.

Figure 12. The behavior of $J(\omega)$ for different correlation times (τ_c) (a) $\tau_c > 1/\omega_n$, (b) $\tau_c \sim 1/\omega_n$ and (c) $\tau_c < 1/\omega_n$ where ω_n is a frequency of interest for which W is calculated. The value of W as a function of τ_c is given in d and plotted as a log/log scale in e (from Neuhaus).



The NOE enhancements where $f_I(S)$ is the fractional change in NOE can be expressed by equation 10.

$$f_I(S) = \frac{(I_z - I^\circ_z)}{I^\circ_z} \qquad 10$$

where S is the saturated spin and I^0_z is the equilibrium intensity of the I spin transition. Through the use of the Solomon equations (Solomon, 1959) the relationship of the NOE to the W rates and the γ_S/γ_I (upon saturation of spin S and observing spin I) is as follows:

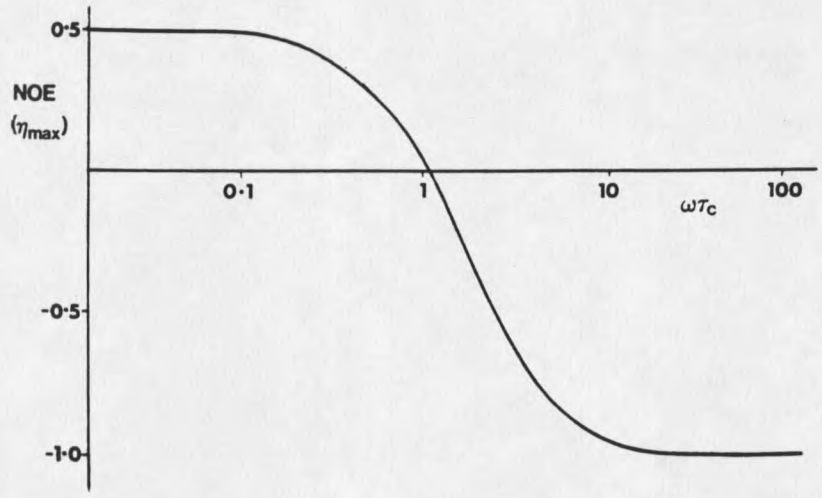
$$f_I(S) = \frac{(I_z - I^\circ_z)}{I^\circ_z} = \frac{\gamma_S}{\gamma_I} \frac{W_{2IS} - W_{0IS}}{W_{0IS} + W_{1IS} + W_{2IS}} \qquad 11$$

$(W_{2IS}-W_{0IS})$ describes the net rate of NOE enhancement which is usually called the cross relaxation σ_{IS}. The term $(W_{0IS}+W_{1IS}+W_{2IS})$ is the total dipolar longitudinal relaxation (ρ_{IS}) (Neuhaus and Williamson, 1989). Equation 11 can be simplified as to

$$f_I(S) = \frac{\gamma_S}{\gamma_I} \frac{\sigma_{IS}}{\rho_{IS}} \qquad 12$$

The relationship between the maximum NOE and the correlation time τ_c can be seen in figure 13.

Figure 13. Dependence of the proton NOE enhancement upon $\omega\tau_c$ (Neuhaus and Williamson, 1989).



As can be seen the maximum NOE intensity enhancement for a small molecule where the W_2 transition dominates is 0.5. While in the slow tumbling region, where W_0 transition dominates, a -1 negative NOE enhancement occurs.

The 1D steady state NOE experiment has been used widely in the past for distance determination in large and small molecules. For 1D NOE experiments selective irradiation of a resonance is used which can be difficult to implement for complex spectra. In a 2D NOESY experiment nonselective excitation is used, and all the resonances are sampled simultaneously by frequency labelling during the variable t_1 time. In addition, spreading of the NMR resonances over two

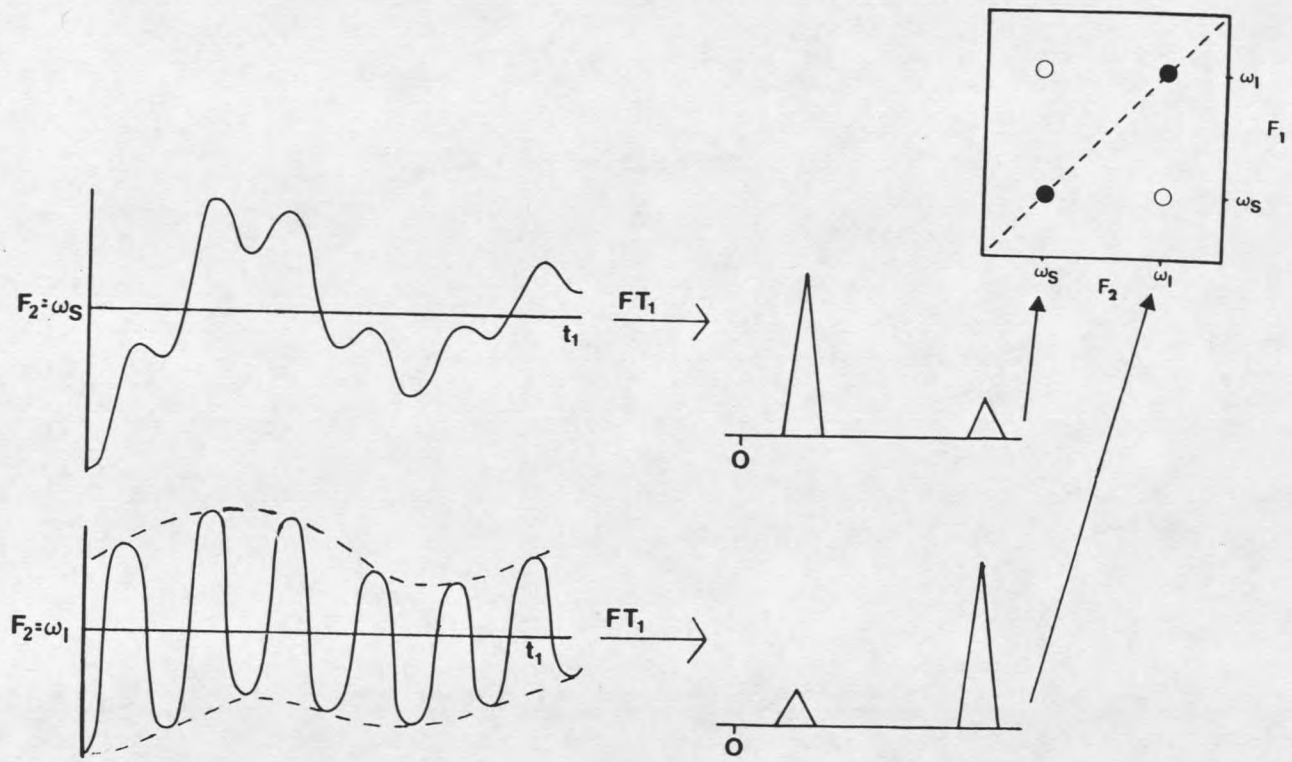
dimensions increases effective spectral resolution. Although 2D NOESY is actually less sensitive than a 1D NOE experiment, advantages in easier experimental setup, a tendency to provide better baselines and far superior selective "irradiation" gives 2D NOESY experiments a definite practical edge for moderately sized biomolecules.

The 2D NOESY consists of three 90° pulses (figure 9). In a simple two spin system; I and S precess in a magnetic field at frequencies of ω_I and ω_S. At the end of the t_1 period spins I and S have precessed through the angles $(\omega_I-\omega_0)t_1$ and $(\omega_S-\omega_0)t_1$ in the rotating frame and the net magnetization have amplitudes of $\cos(\omega_I-\omega_0)$ and $\cos(\omega_S-\omega_0)$ along y. A second $90°_x$ pulse rotates the y components of the I and S magnetization to the z axis. The amplitudes of the spins are now frequency labelled as a function of the t_1 delay and the z component of the magnetization of spin S and I are $\cos(\omega_S-\omega_0)t_1$ and $-\cos(\omega_I-\omega_0)t_1$ respectively. After the second pulse cross relaxation occurs during the mixing time (τ_m) and the frequency of the I vector $-\cos(\omega_I-\omega_0)t_1$ is amplitude modulated at the frequency of the S vector $-\cos(\omega_S-\omega_0)t_1$ by cross relaxation (σ_{IS}). The S vector also experiences a modulation of its amplitude $-\cos(\omega_S-\omega_0)t_1$ at the frequencey of the I spin $-\cos(\omega_I-\omega_0)t_1$ by cross relaxation. The effect of the frequency modulations of both spins can be seen in figure 14. The magnitude of the

amplitude modulation of I and S frequencies depends on the size and sign of the NOE enhancement during the mixing time which in turn depends on $1/r^6_{IS}$. The third 90° pulse puts the amplitude modulated spin vectors of I and S in the transverse plane for detection during t_2. The result of spins I and S frequency modulations upon one another when Fourier transformed (figure 14) results in crosspeaks in the 2D spectra which are correlated to one another as schematically shown in figure 14 (Neuhaus and Williamson, 1989).

By measuring the intensities of the crosspeaks and the cross relaxation rate σ_{IS}, the NOE enhancement can be attained. Unfortunately extraction of the pure cross relaxation rate is not so easy. Cross relaxation, σ_{IS}, at short mixing times is linear with τ_m but becomes nonlinear at longer mix times. At short mix times there is a weak signal and a poor signal/noise ratio. The nonlinearity at higher mix times is due to T_1 relaxation along z and spin diffusion where magnetization may be lost from spins I and S by way of other spins. The effect of T_1 can be compensated by dividing the cross peak volumes by diagonal peak volume to obtain fractional NOEs. The internuclear distance from a NOESY crosspeak is usually calculated by comparison of the intensity to a cross peak from a pair of protons at a known distance at sufficiently short mixing time.

Figure 14. Representation of a NOESY experiment to show the origin of crosspeaks. On the left is the F_2 frequency for spins I and S; and their effect of modulation on the frequency and resulting FT. The right shows the resulting diagonal and correlated crosspeaks in a 2D spectra (Neuhaus and Williamson, 1989).



The difficulty with this is that spins at different distances have different disturbances by spin diffusion. The cross relaxation in NOESY is approximately proportional to $1/r_{ij}^6$ and a known calibration distance from protons on an aromatic ring, a β methylene, or sidechain amine, can be used: $d_{ij}=d_{cal}(\sigma_{cal}/\sigma_{ij})^{1/6}$. This method of distance calculation is known as the isolated two spin approximation (ISPA).

The ISPA approach for distance determination is prone to errors due to conflicting effects of low signal to noise ratios at short mixing times, and spin diffusion as well as baseline artifacts. In order to determine more accurate distances from 2D NOESYs, methods have been developed for a complete relaxation matrix analysis to correct for spin diffusion effects and to allow use of higher signal to noise data at longer mixing times. Full relaxation matrix analysis methods such as IRMA (Boelens et al., 1989); CORMA (Borgias et al., 1988); and MARDIGRAS (Borgias et al., 1990) use the experimental NOESY intensities, and an initial approximate structure to take account of the effect of nearby protons and calculate a set of more accurate distances. These complete relaxation analysis methods are usually iterated. Ideally as the structure converges to a more accurate structure more exact corrections for spin diffusion effects can be accomplished.

Cross Relaxation Theory

The relaxation behavior of a system of protons can be

described as:

$$\Delta M_z = R\Delta M_z \qquad 13$$

where ΔM_z is the deviation of the magnetization from equlibrium, and R is the relaxation matrix which is depicted as:

$$R = \begin{bmatrix} R_{ii} & R_{ij} & . \\ . & . & . \\ . & . & . \end{bmatrix} \qquad 14$$

where the off diagonal terms (R_{ij}) represent cross relaxation between two spins (i&j) and the diagonal terms (R_{ii}) represent the sum of all relaxation pathways.

$$R_{ii} = \Sigma (W_0^{ij} + 2W_1^{ij} + W_2^{ij}) + R_1 i \qquad 15$$

$$R_{ij} = (W_2^{ij} - W_0^{ij}) \qquad 16$$

As noted before the term (W_2^{ij}-W_0^{ij}) describes the net rate of transitions that change spin label populations and which give rise to NOE enhancements. The term R_{1i} takes into account external sources of relaxation such as paramagnetic impurities or binding to large molecules. The rates of the different W relaxation transitions can be related to the correlation time, resonance frequency, and internuclear distances.

$$W_0^{ij} = q\frac{\tau_c}{r_{ij}^6} \qquad 17$$

$$W_1^{ij} = 1.5\frac{q\tau_c}{r_{ij}^6}\frac{1}{1+(\omega\tau_c)^2} \qquad 18$$

$$W_2^{ij} = 6\frac{q\tau_c}{r_{ij}^6}\frac{1}{1+4(\omega\tau_c)^2} \qquad 19$$

where $q=0.1\gamma^4\hbar^2$. The solution of the system of equations 13-19 can be setup as such:

$$M\tau_m = a(\tau_m)M(0) = e^{-R\tau_m}M(0) \qquad 20$$

The a is the matrix of mixing coefficients which are proportional to the 2D NOE intensities and M and M(0) are magnetization vectors at t and at equilibrium respectively. The rate matrix R describes the dipole-dipole interactions. The evaluation of intensities from the a matrix can be analyzed in an isolated spin pair approach (ISPA) approach using a Taylor series expansion:

$$a(\tau_m) = e^{-R\tau_m} \approx 1 - R\tau_m + \frac{1}{2}R^2\tau_m^2 - \ldots + \frac{(-1)^n}{n!}R^n\tau_m^n + \ldots \qquad 21a$$

This expression can be truncated after the first linear R

term which simplifies the analysis as long as the mixing time (τ_m) is kept reasonably short. This approach however suffers from the neglect of network relaxation and assumption that spins are isolated. A solution to this problem is to generate a complete matrix from a combination of experimental data and calculation of the missing elements. The a coefficients taken from the experimental NOE intensities where available and the missing elements are evaluated using proton-proton distances from an initial trial structure (Keepers et al., 1984) (Baleja et al., 1990). The complete a matrix can be solved by using a numerical approach. The R matrix is recast: $R=\chi\lambda\chi^{-1}$, where χ is the eigenvector matrix and λ is the diagonal matrix of eigenvalues. The a matrix can then be analyzed as follows:

$$a = 1 - \chi\lambda\chi\tau_m + \frac{1}{2}\chi\lambda\chi^{-1}\chi\lambda\chi^{-1}\tau_m^2 - \ldots. \qquad 21b$$

$$a = \chi e^{-\lambda\tau_m}\chi^{-1} \qquad 21c$$

So a refinement program such as MARDIGRAS can calculate the exponential by diagonalizing the R matrix and then equation 21c can be evaluated to obtain theoretical NOE intensities. The method takes into account spin diffusion effects on cross peak intensities which is not addressed by the ISPA approach.

The refined internuclear distances from complete

relaxation analysis can be used in the calculation of structures using molecular dynamics or simulated annealing. Molecular dynamics and simulated annealing use Newton's equation:

$$F_i = m_i a_i \qquad 22$$

where F_i is the force, m_i is the mass and a_i is the acceleration of particle i. The force on a system can be computed from the derivative of the potential energy (V) with respect to coordinates r_i and equation 22 can be expressed as:

$$F_i = \frac{-\delta V}{\delta r_i} = m_i \frac{\delta^2 r_i}{\delta^2 t} \qquad 23$$

The potential energy of the system can be approximated using a series of energy terms to account for the effects of distances, bond angles, and electrostatic interactions. The collection of terms that describe the potential energy of a molecule as a function of the above variables is often called the molecular forcefield. The individual energy parameters which make up the forcefield can be seen below:

$$F_{Tot} = F_{nonbond} + F_{cov} + F_{NOE} \qquad 24$$

where F_{cov} are the terms to describe the energetic effects of changes in bond lengths, angles, chirality and planes.

$$F_{cov} = \sum_{bond} K_b(r-r_0)^2 + \sum_{ang} K_\theta(\theta-\theta_0)^2 + \sum_{im} K_\theta(\theta-\theta_0)^2 \qquad 25$$

$$+ \Sigma_w K_w(1+\cos\omega) \qquad 25$$

Nonbonded interactions (Van der Waals) can be represented by a Lennard Jones term:

$$F_{nonbond} = \sum \epsilon[(r^*/r)^{12} - 2(r^*/r)^6]$$

where r is the Van der Waals radius and ε is the dielectric constant. The terms that utilize NMR distance restraints, F_{NOE}, can be represented by:

$$F_{NOE} = \begin{cases} K_{NOE}\ (r_{ij}-r_{ij}^u)^2 & \text{if } r_{ij} > r_{ij}^u \\ 0 & r_{ij}^l < r_{ij} < r_{ij}^u \\ K_{NOE}\ (r_{ij}-r_{ij}^l)^2 & r_{ij} < r_{ij}^l \end{cases}$$

where r_{ij}^u and r_{ij}^l are upper and lower limits of the distances, and K_{NOE} is the force constant.

With an expression for the potential energy, and the known masses of the particles in the system, it is possible to solve equation 22 for a one or two particle system. For more complicated systems numerical methods are needed to solve the motions of a multiparticle system. A Taylor series can be used to express the motions of a multiparticle system:

$$r(t+\Delta t) = r(t) + \frac{\delta r}{\delta t}\Delta t + \frac{\delta^2 r}{\delta t^2}\frac{\Delta t^2}{2} + \ldots \qquad 26$$

This equation depends on knowing the initial position r(t), velocity ($\delta r/\delta t$), and acceleration ($\delta^2 r^2/\delta t^2$) of each particle. The coordinates of position can be taken from an initial structure. The initial coordinates can either be taken from an x-ray structure (if available), calculated from a protocol called simulated annealing (SA) or generated in a variety of other ways.

SA can start with atoms distributed randomly in space. The force constants of the target function in the forcefield are initially set low and temperature and kinetic energy (T) set high. This will randomize atoms in the beginning and the high temperatures and low force constants will allow the system to explore conformational space and break out of false minima. A series of steps of restrained molecular dynamics (RMD) is done while the force constants of F_{NOE}, F_{cov}, F_{repel} are incrementally increased and temperature decreased in stages.

The initial velocity for equation 26 can be obtained from a Maxwell-Boltzmann distribution for the temperature of interest. The accelerations can be calculated from equation 23. The term $r(t+\Delta t)$ movements of the structure is found in equation 26. The original coordinates are replaced by the new coordinates, velocities are updated by subtracting the new and old coordinates and dividing by a time step, and the

acceleration is updated by computing the gradients with the new coordinates (Biosym, 1993) (Nilges et al., 1988). Once this cycle is complete, it can be repeated many times. Simulated annealing and molecular dynamics allow for searching of conformations in space of a system of atoms. Inclusion of NMR derived distance constraints allow the NMR information to guide the SA and MD toward the 3D structure.

Modified EF-hand/Bacteriorhodopsin Background

We have investigated the 3D structure of a model for one of the cytoplasmic surface loops of the membrane protein (BR). The cryoelectron microscopy method on tilted sections has been used to study the structure of intact BR in membranes. This approach provides the most information about the seven transmembrane helical regions in the membrane but provides rather little information on the surface loops connecting the helices in the aqueous regions.

There is experimental evidence for functionally important cation binding to BR. The native membrane is purple and absorbs maximally at λ_{max}=568 nm in its light adapted state. When cations are stripped from the purple membrane the color changes to blue (maximum absorbance λ_{max}=608 nm) and the ability to pump protons is lost (Chang et al., 1985)(Chang et al., 1986)(Jonas et al., 1991). The regeneration of the purple color and proton pumping can be restored by a number of different cations such as Gd^{+3}, Eu^{+3}, La^{+3}, Cr^{+3} Cd^{+2}, Ba^{+2}, and Ca^{+2}. The highest charge density

ions have the highest affinity for binding sites in bacteriorhodopsin (BR), and restore the purple color and proton pumping at the lowest metal ion concentrations (Ariki et al., 1986).

It was proposed some time ago that a cation binding site may be located in the loop connectiong C and D helices of BR (Helgerson et al., 1988). A transmembrane folding model of BR is shown in figure 16 which indicates the seven helices designated A-G and the connecting loops using the one letter amino acid code that is explained on the following page. Residues 102Asp, 104Asp, 105Gln, 107Thr, were proposed as part of a cation binding site. This sequence has homology with the EF-hand domains of calcium binding proteins, such as calmodulin with one deletion (Helgerson et al., 1988) (table 6). However, in order to be homologous with EF-hand structures a metal coordination site is "missing" from the above group of residues and this was proposed to be filled by 161Glu or 166Glu that are located on the EF loop of BR. The use of a remote Glu to complete the metal coordination in an EF-hand is known in the structure of a calcium binding site of an E.Coli galactose binding protein. In the galactose binding protein a remote Glu about 60 residues away in the sequence completes an EF-hand binding site (Vyas et al., 1987). The one letter amino code is: A=Ala, C=Cys, D=Asp, E=Glu, F=Phe, G=Gly, H=His, I=Ile, K=Lys, L=Leu, M=Met, N=Asn, P=Pro, Q=Gln, R=Arg,

S=Ser, T=Thr, V=Val, W=Trp, and Y=Tyr.

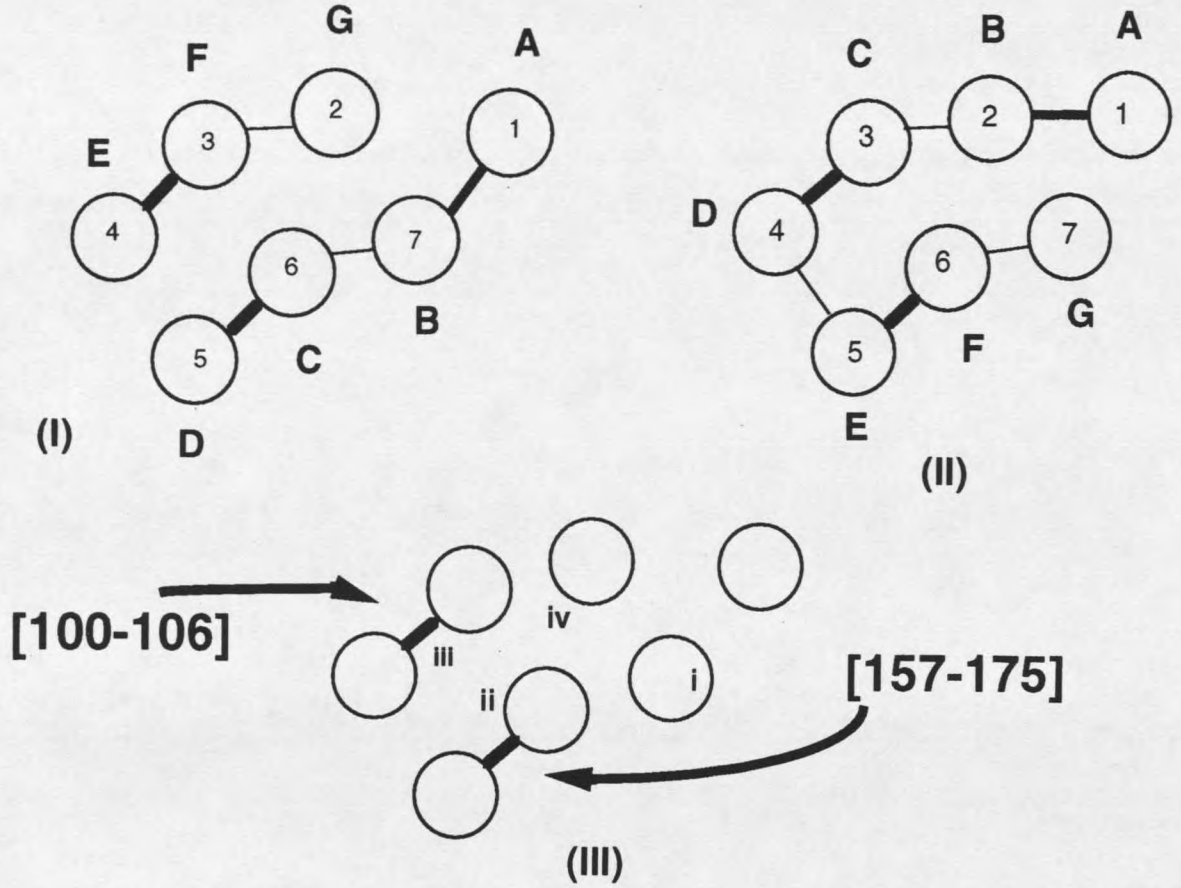
The classic EF-hand Ca^{+2} binding motif contains 12 amino acid residues flanked by 2 helices. For the most part, five residues are involved in providing oxygen ligands to the metal ion, and the other residues help to stabilize the coordination geometry. The majority of the EF-hand Ca^{+2} binding sites are seven coordinate with a oxygen to metal ion distance of about 2.4 Å. The coordination of the metal ion is not always entirely due to amino acid residues on the protein since water may also be involved in the coordination site. The seven bond coordination often appears to form a pentagonal bipyramidal geometry in typical EF-hands (Strynadka et al., 1989).

Evidence has been reported for several sites of cation binding on BR. X-ray diffraction difference density maps of membranes reconstituted with lead compared to those reconstituted with calcium show (Pb-Ca) density near helix 7(G), the carboxy terminus. A second binding region is located between helix 5(E) and 6(F), and a third between helix 3(C) and 4(D) (figure 15) (Katre et al., 1986). The addition of $[Co(NH_3)5H_2O]^{+3}$ to blue membranes regenerated the photocycle. Inner sphere cobalt (III) complexes change ligands slowly and were used to locate Co^{+3} binding sites in the BR sequence. The Co^{+3} loaded membrane was enzymatically cleaved and Co^{+3} was bound to two peptide fragments containing sequences

Table 6. Primary sequence homologies comparing known calcium-binding sites to sites of interest in BR and GPIIb-IIIa. Bold faced residues emphasize homologous regions.

Protein Sequence	Possible calcium-liand residues																
			(x)		(y)		(z)		(-y)		(-x)				(-z)		
Parvalbumin (88-104)	A	G	**D**	S	**D**	G	D	**G**	K	**I**	G	V	D	E	F	T	A
Calmodulin (48-75)	E	V	**D**	A	**D**	G	N	**G**	T	**I**	D	F	P	E	F	L	T
Bacteriorhodopsin (100-115)	**L**	V	**D**	A	**D**		Q	**G**	T	**I**	L	A	**L**	**V**	**G**	**A**	D
Glycoprotein IIB																	
(272-288)	V	G	E	F	**D**	G	D	L	N	T	T	E	Y	**V**	**G**	**A**	P
(326-342)	V	T	**D**	V	N	G	D	**G**	R	H	D	L	**L**	**V**	**G**	**A**	P
(393-409)	**L**	G	**D**	L	**D**	R	D	**G**	Y	N	D	I	A	**V**	A	**A**	P
(455-471)	A	V	**D**	I	**D**	D	N	**G**	Y	P	D	L	**I**	**V**	**G**	**A**	Y
Galactose-Binding Protein (132-148)	G	W	**D**	L	N	K	D	**G**	Q	**I**	Q	F	V	L	L	K	G

Figure 15. I and II models assigning transmembrane helices from the sequence folding model (A-G) to the helical structures seen by three dimensional electron microscopic structural analysis. The helix linker regions are shown from the cytoplasmic ( ) and periplasmic (--) sides of the membrane. (III) location of cation binding sites (i-iv) observed using X-ray diffraction difference density maps.



102(-Asp-Ala-Asp-) and 232(-Glu-Ala-Glu) both of which have carboxyl groups favorable for binding Co^{+3}. These experiments indicate metal ion binding sites are present between helix (C) and (D) and on the C-terminal tail in BR (Engelhard et al., 1989).

La^{+3} was used to assess the metal binding affinity in the peptides of interest in solution by observing binding-induced shifts in ^1H NMR. The trivalent La^{+3} has a higher affinity than divalent Ca^{+2} and does not produce paramagnetic shifts or substantial broadening of the ^1H NMR resonances. Marsden and Hodges showed that by decreasing the number of Asp residues in the (X), (Y), and (Z) positions of EF-hand peptides, the La^{+3} ion affinity decreased in solution. Therefore it is seen by site specific mutagenesis that alteration of key residues of and EF-hand sequence effect metal ion affinity.

The BR sequence between (100-115) has residues in the (X), (Y), (Z) and (-Y) positions with sequence homology to that of the EF-hand sequence in parvalbumin and calmodulin (Table 1). However the BR sequence has a deletion and substitution compared to a consensus EF-hand sequence. The Gly at position (Z-1) is missing and the Asn at (Z) in calmodulin is replaced by Gln in the BR loop (table 6). In a typical EF-hand an invariant Glu residue is found at the (-Z) position of the 12 residue sequence. In BR(100-115) the (-Z) to (-Z+2) would be occupied by Val-Gly-Ala if the

sequence was truely homologous to an EF-hand (Strynadka et al., 1989).

The x-ray structure of E. Coli galactose binding protein shows that the (-Z) calcium binding position is not in the twelfth position of the loop domain (Vyas et al., 1987). In this protein there is a nine residue sequence (134-142), which is homologous in sequence and structure to the (X) to (-X) region of an EF-hand domain. As in the GPIIb and BR(100-115) sequences, the galactose binding protein does not have an invariant Glu in position (-Z), but has a Leu (table 6). The Glu that occupies the (-Z) position in the binding site in the galactose binding region is 60 residues away in the primary sequence at position 205. Even though this particular Glu is 60 residues away, it is in the proper position in the tertiary structure of the galactose binding protein, it serves as the (-Z) calcium ligand in the crystal structure. The calcium binding protein GPIIb has four partial EF-hand sequences that are homologous to calmodulin but are also missing the (-Z) Glu in the local 12 amino acid peptide in an analogous manner to the E.Coli galactose binding protein and BR.

In BR there are two Glu residues (161,166) on the EF loop that are candidates to fill the (-Z) Ca^{+2} coordination. The two Glu residues are approximately 60 residues away from the major part of the potential cation binding sequence (102-109) which is in the linking region between

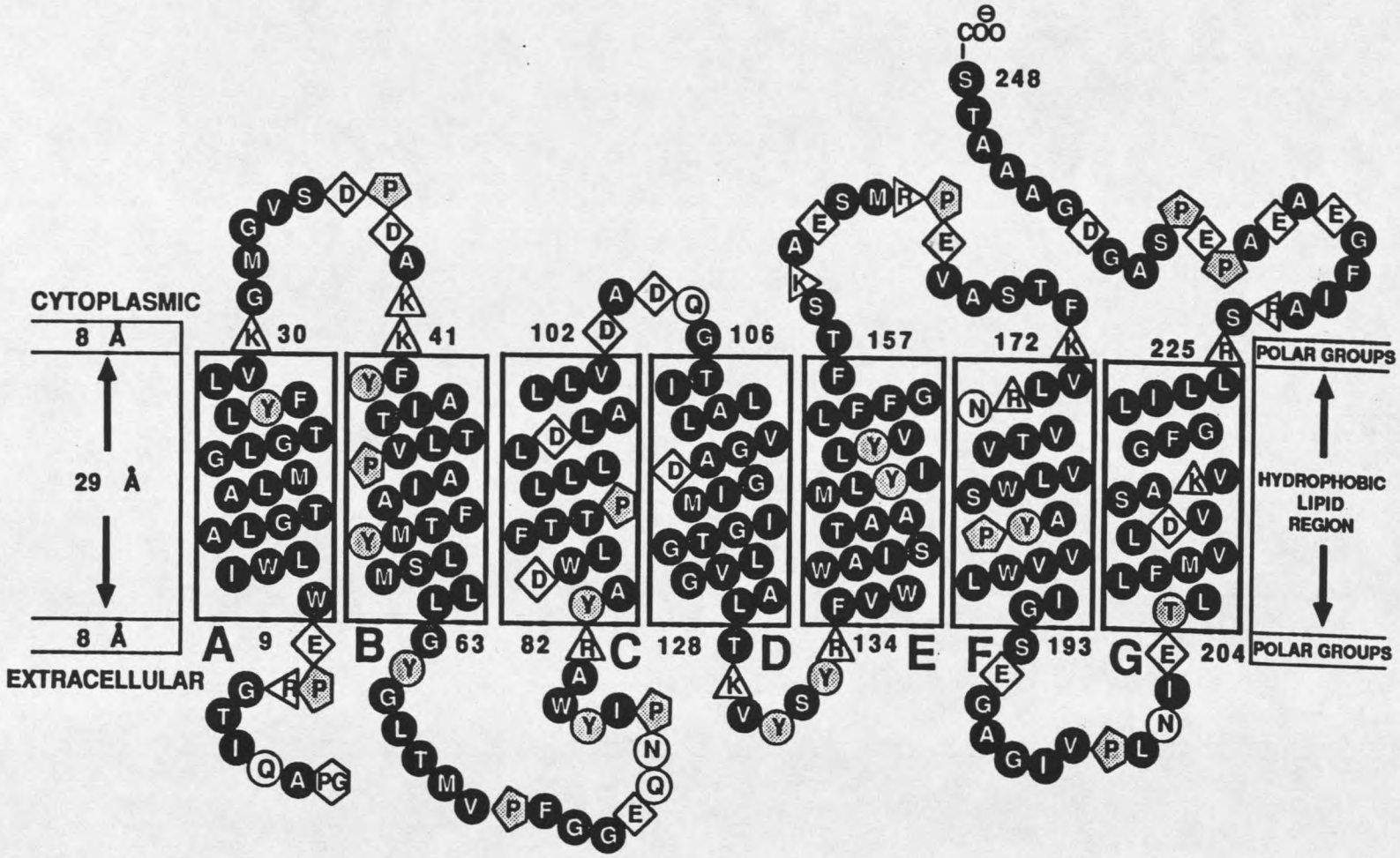
transmembrane helices C and D. The two Glu residues of interest at (161 and 166) are on the same side of the EF connecting loop on the membrane as the C-D loop (figure 16). There is evidence which indicates that these two helix linking regions are close to each other in the BR structure (Agard et al., 1982) and are also in close proximity to cation binding regions determined in X-ray diffraction difference density maps (figure 15)(Katre et al., 1986).

In summary the 102-109 region of BR is homologous to the (X) to (-X) of a normal EF-hand sequence. Comparing this region to the calmodulin site (48-75), the Gly at position (Z-1) is missing and the Asn at (Z) is replaced by Gln (table 6). The missing Glu at position -Z may be filled by a distant (161 or 166) Glu by analogy to the E.Coli galactose binding protein.

Materials and Methods

A preliminary test of the plausibility of metal ion binding by the modified EF-hand sequence in the CD loop of BR was carried out. The parvalbumin crystal structure in the Brookhaven protein data bank was first energy minimized using Amber (Weiner et al., 1981). The parvalbumin third EF-hand in the energy minimized structure was then modified with the deletion of Gly and substitution of Asp (Z) with Gln. The modified protein with Ca^{+2} bound was then energy minimized with Amber.

Figure 16. The seven transmembrane model of Bacteriorhodopsin showing the loop between helix C and D, Glu-161 and Glu-166 (shaded). The one letter code for the amino acids is explained on page 66 of the text.



The cation ion binding site geometry was altered substantially but a seven coordinate tertiary metal binding structure remained with six carboxyl oxygens to the metal ion which will be discussed more fully in the results.

Peptides with free carboxyl terminal residues were synthesized using a Milligen Model 9050 continuous flow automated peptide synthesizer with F-moc chemistry. The coupling cycle time was extended 50% for the aspartate residues to increase coupling yields. F-moc-phenylalanine-PepSyn-KA resin (Milligen/Biosearch) was used to initiate the synthesis and the completed peptides were cleaved from the resin using 95% TFA/5% phenol. The crude peptides were isolated by ether precipitation. The purity of the peptides was determined to be > 90% by reverse phase HPLC and mass spectrometry. The amino acid composition of the peptides was determined by HPLC analysis of phenylisothiocarbamyl derivatives after acid hydrolysis of the peptides in vacuo (Meredith et al., 1984). The amino compositions obtained agreed with the sequences.

The NMR spectra of the Parv(89-102) peptide and the modified Parvalbumin peptide (GDSDQGKIGVDEF) with added metal were recorded on 2.5 mM samples dissolved in 400 ul of 20 mM deutero imidazole and 100 mM KCl and 100 ul of D_2O adjusted to a pH of 6.35. The peptide samples were titrated with small aliquots of 1M $LaCl_3$ solution and their NMR spectra measured. Both samples were titrated to a final

peptide-metal ion ratio of 1:3 peptide/La^{+3}.

NMR spectra were measured using a Bruker AM 500 spectrometer with sequential quadrature detection. TOCSY spectra (Braunschweiler et al., 1983) were collected at 1° C with a MLEV-17 spin lock time of 50 ms (Bax et al., 1985) NOESY spectras (Jeener et al., 1979) were obtained in the temperature range of 1-30 °C with a 400 msec mixing time using jump and return for solvent suppression (Plateau et al., 1982). All 1D spectra were collected in a temperature range of 1-30°C.

Two-dimensional NMR spectra were processed on SGI Iris workstations using FELIX macro programs developed in our lab by Christophe Lambert. Quadrature detection in t_1 was achieved by TPPI (Marion et al., 1983) for all 2D experiments. Apodization functions applied to 2D spectra were 60 degree shifted sine bell window in t_1 and t_2 with zero filling to 2048 complex points. Spectral widths of 5263 Hz were used for all spectra. Polynomial baseline corrections of the third order were used in f_2 for both NOESY and TOCSY. The preacquisition delay and transmitter phase were finely adjusted to avoid any rephasing of the spectra and therefore improve baseline flattening (Marion et al., 1988). The NOESY spectra used 1024 points in t_1 and the D0 delay in t_1 was adjusted to be equal to IN/2=47.5 usec minus a $4\tau_{90}/\pi$ to account for the finite hard pulse length (Bax et al., 1991). The phase for the t_1 to f_1

transform was PH0=45 and PH1=-90 (Bax et al., 1991). A correction was also applied to the NOESY intensity in f_2 before the f_1 to compensate for the intensity distortion due to the jump and return excitation and to retrieve intensity loss especially in the alpha proton region close to water. Peak volumes were obtained by manually selecting boxes around crosspeaks and summing the points in the box.

NOESY cross peaks were converted to approximate proton distances using the two spin approximation, using sidechain amino protons of Asn as a reference distance (1.8 Å). A total of 116 NOE cross-peaks were identified in which there were 41 i,i+1 and 8 i,i+2 interresidue constraints. The NOESY distances were converted to a distance restraint file using the FELIX database software package. The stereospecificity of β methylenes were not assigned initially and pseudo atom corrections as described (Wuthrich, 1983) were applied to the distance restraints. After pseudo atom corrections 95 NOE distances were used for the initial approximate structure calculations. The NOE distance constraints were organized in three classes. Initially NOE distances less than 2.5 Å had an upper limit constraint limit of 2.5 Å and a lower limit of 1.8 Å. NOE distances less than 3.5 Å but greater than 2.5 Å had a 0.5 Å error range added to create lower and upper limits. NOE distances above 3.5 Å had a lower limit of 3.5 Å and no upper limit error range. All NOE distance constraints had

force constants of 100 kcal/(mol/Å^2). Chirality restraints were applied on all chiral atoms. Restraint on the omega dihedral angles was applied with force constant of 300 kcal/(mol/rad^2) to maintain trans peptide configurations. Molecular dynamics, and simulated annealing calculations were in the Discover v2.9 package and a variety of analysis routines in the InsightII v2.2 and NMRchitect v2.0 packages (Biosym, 1993).

Solution structures of the lanthanum complex of the modified EF-hand peptide were studied by using NMR distance constraints initially using the two spin approximation for the restraints as described above, along with seven metal to ligand restraints. These oxygen to metal distance restraints were obtained from an AMBER minimization of the modified metal ion binding site in parvalbumin as explained earlier. All molecular dynamics was carried out with the CVFF forcefield using direct velocity scaling and temperature bath coupling (Dauber et al., 1988). The potential of the metal atom used was that of Ca^{+2} since at that time a potential for La^{+3} was not available.

The approximate structure obtained was employed to improve the NOE distances and to obtain estimates of distance errors using MARDIGRAS complete relaxation matrix analysis (Borgias et al., 1990). The Mardigras derived distance constraints and distance errors were used in a simulated annealing protocol (Biosym, 1993) involved 60,000

steps of dynamics from 800°K and cooling to 300°K ending with four rounds of steepest descent and by four rounds of conjugate gradient minimization.

Results

At the time this study was carried out the atomic coordinates for calmodulin were not available. Therefore the effect of the Gly deletion and Asn to Gln modification on the EF-hand ion binding structure was investigated using the third EF-hand (90-101) in the known x-ray structure of parvalbumin as a model.

The crystallographic structure of parvalbumin was used to investigate the probable consequences of an amino acid deletion and substitution in an EF-hand site. First the parvalbumin crystal structure was energy minimized using AMBER (Weiner et al., 1981) to provide the starting structure. The 93Gly residue was removed and the 94Asp was changed to Gln. The resulting approximate coordinates were used as input parameters for energy minimization using AMBER. The energy minimized structure of the modified Ca^{+2} complex of the protein indicated that the geometry of the cation binding site was substantially modified but retained a tertiary binding structure with suitable bond lengths to the calcium ion as shown in table 7. In the modified parvalbumin, both carboxylate oxygens of 90Asp coordinate to the metal ion, while 93Gln and the water molecule at the (-X) position are no longer involved in coordination with the

calcium ion. The distances between the calcium ion and oxygen atoms found for coordination of the modified and normal EF-hand of parvalbumin are shown in table 7.

Table 7. Calcium-oxygen distances for the calcium-binding residues of native and modified parvalbumin.

			Distances (Å)	
Residue	Group		Native	Modified
Asp 90	Carboxylate	OD1	------	2.28
		OD2	2.23	2.29
Asp 92	Carboxylate	OD1	2.31	2.34
		OD2	2.24	2.27
Asp 94 (Gln 93)	Carboxylate	OD1	2.51	------
		OD2	-----	------
Lys 96 (95)	Carbonyl		2.33	2.34
Glu 101	Carboxylate	OE1	2.29	2.29
			2.38	2.29

The putative structure of the modified parvalbumin is certainly not to be considered a proof of the metal ion binding capability or geometry of the site of interest in BR, but the results were encouraging and further experiments were carried out.

We studied the metal ion binding affinity of the parvalbumin (90-101) in solution and the modified (90-101) parvalbumin with deletion of 93Gly and D94Q mutation using

1D NMR. La^{+3} has been a useful tool in analogous studies of the ion binding capabilities of modified EF-hand peptides (Marsden et al., 1988). Trivalent La^{+3} is used instead of divalent Ca^{+2} due to its stronger complex formation and its diamagnetic NMR properties. Due to the high positive charge density, La^{+3} a very close match in atomic radius to that of Ca^{+2} (Weast, 1986). As can be seen from figure 17 a and b, several resonances in both the normal and modified parvalbumin sequences respectively shift considerably upon addition of La^{+3}. The Gly alpha resonances show the greatest shift in the NMR spectra upon addition of La^{+3} as has been observed in other EF-hand analog peptides in solution (Marsden et al., 1988). Because of overlap in other resonances, the Gly alpha shifts were used to follow the La^{+3} titrations. The perturbation of the chemical shifts result from conformational and environmental changes due to metal ion binding since La^{+3} is diamagnetic. It can be seen from figure 18 the normal parvalbumin sequence showed tight binding and reached a maximal chemical shift at La^{+3}/peptide ratio of 1:1. The NMR resonances of the modified parvalbumin sequence continued to shift upon addition of La^{+3} indicating a lower binding affinity.

The La^{+3} binding studies by the peptides in solution as well as the initial energy minimizations on the modified EF-hand segment of parvalbumin indicate that the modified EF-hand sequence can bind cations but with a lower affinity and

Figure 17. a. 1D NMR of the aliphatic region of the normal parvalbumin (89-102) sequence; lines indicate Gly-αH ^{1}H resonance (lower), and the normal parvalbumin (89-102) titrated with La^{+3} (top).

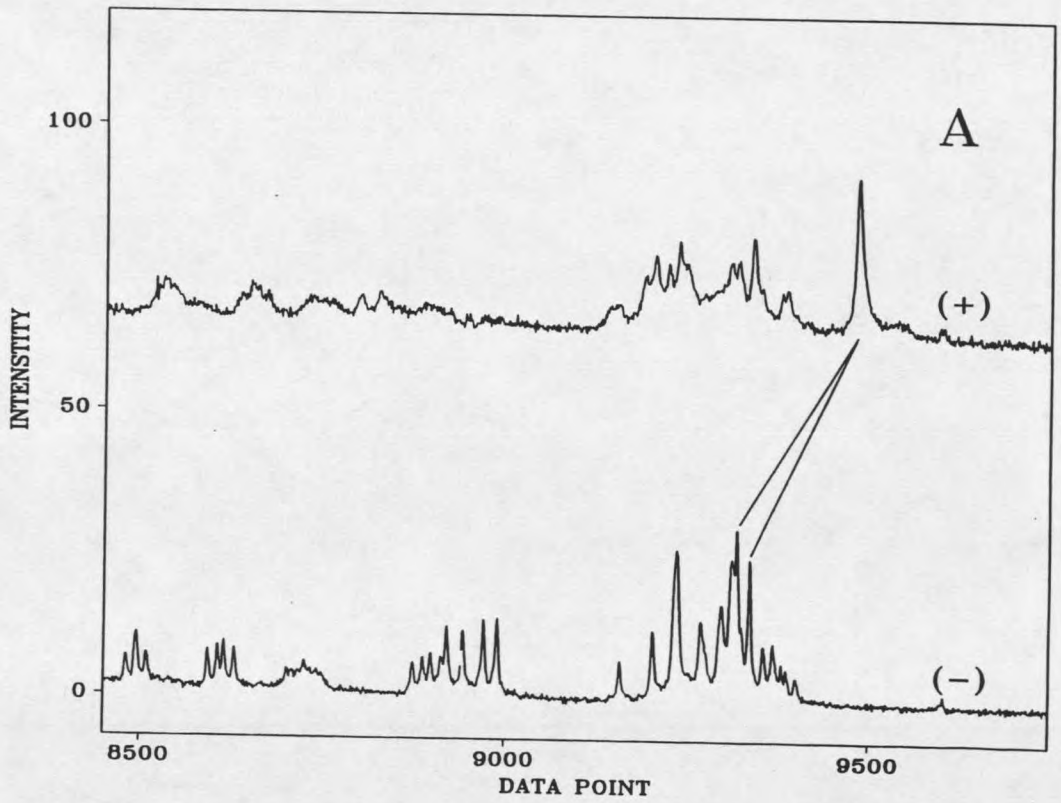


Figure 17 (continued). b. 1D NMR of the aliphatic region of the modified parvalbumin sequence; lines indicate Gly-α ^{1}H resonances (bottom), and the modified parvalbumin sequence titrated with La^{+3} (top).

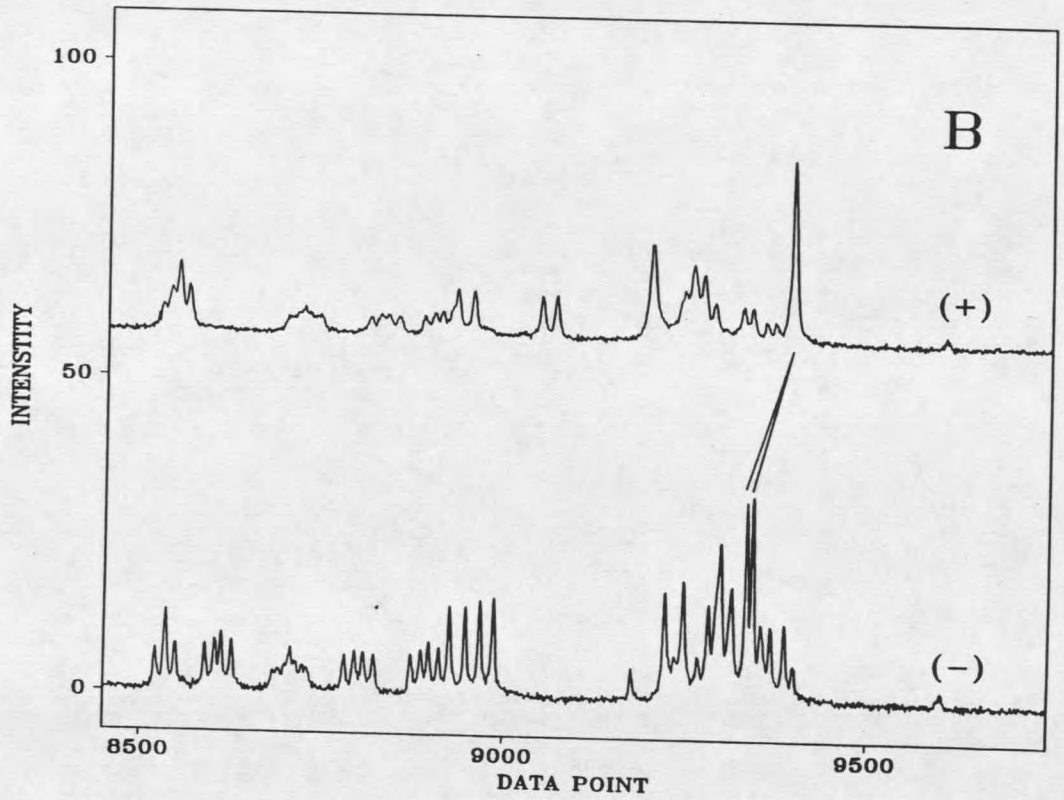
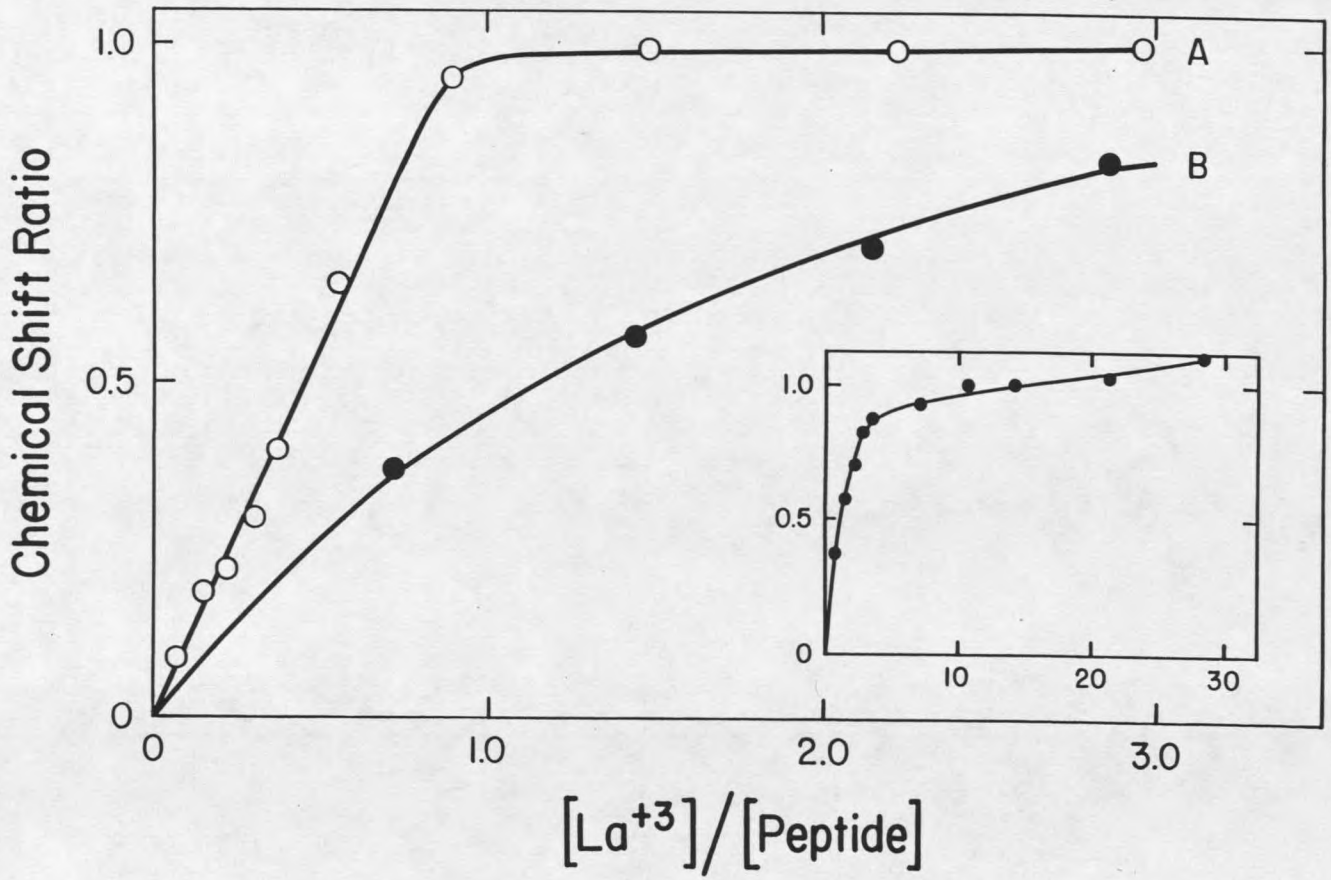


Figure 18. Lanthanum titration of normal parvalbumin (89-102) sequence (A) and modified parvalbumin sequence (B). The plot is for the Gly-α ^{1}H resonances shown in figure 3a and b. The observed chemical shift ratio (CSR) has been normalized for a La^{+3}/peptide ratio of 1:1 for the normal parvalbumin peptide.



probably with a change in geometry. This change in geometry may be responsible for the large (ca. 100x) decrease in metal ion affinity found in the modified sequence (figure 18). The change in geometry and the decreased affinity may also be related to the low metal ion specificity found in the blue to purple transition in BR. This evidence is consistent with the BR (101-113) sequence binding a metal ion with decreased metal ion affinity and specificity compared to an EF-hand because of a distorted EF-hand binding geometry. Therefore, the modified EF-hand sequence of parvalbumin with its homology to BR (102-108) appears to be a plausible starting point for modelling the putative metal ion binding loop for BR.

The assignment of the ^{1}H NMR of the La^{+3} complex of the modified EF-hand peptide in the presence of three fold excess La^{+3} was carried out by 2D NMR. TOCSY spectra were used to establish through-bond connectivity, for side chain spin systems and the NOESY spectra were used to establish interresidue through-space connectivity of spin systems. Certain residues were easily identified by their crosspeak patterns and chemical shifts including 3Ser, 8Ile, 7Lys, 13Phe and 10Val. Even though the peptide contained only 13 residues, the repetition of Asp and Gly residues caused considerable overlap in the NMR spectra. Five backbone amide protons from 12Glu, 5Gln, 8Ile, 2Asp and 4Asp are all within 0.02 ppm of each other but these could be seperated

on the basis of their αH shifts as shown in figure 19. The overlap of the amide backbone protons for the 2Asp and 4Asp in the NOESY (figure 19) makes this assignment ambiguous. Variations in temperature from 1°C to 30°C were carried out to shift amide resonances in the NMR and aid in spectral assignment (data not shown). The data at other temperatures was able to confirm all the assignments but the 2Asp and 4Asp could not be distinguished. Experiments using ^{15}N HMQC would be needed to gain this separation but our NMR instrument was not configured for ^{15}N decoupling so these experiments were not performed. Lorentz-Gauss multiplication apodization was used in 1D and 2D NMR processing to enhance resolution and assist in separation of overlapped amide resonances. The chemical shifts of the ^{1}H NMR assignments which made at 1° C are shown in table 8.

Discussion

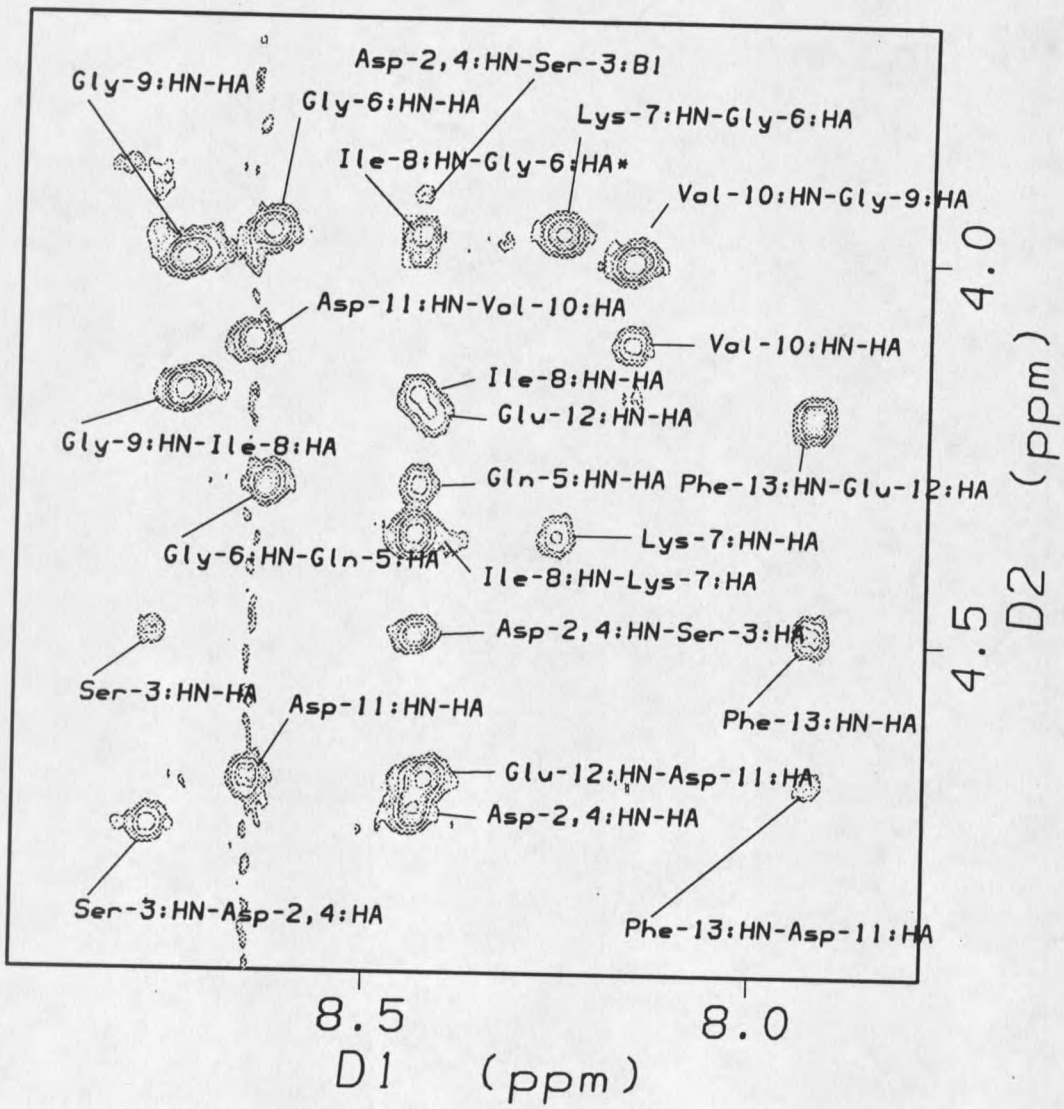
Site specific mutagenesis studies have been done on synthetic peptides (Marsden et al., 1988)(Reid, 1990) (Reid et al., 1981) in order to better understand the basis of the affinity of cation binding to EF-hand loops. By specifically changing certain residues in a peptide, metal ion binding affinities may vary significantly. Marsden and Hodges reported extensive work in which they substituted residues in the (X), (Y), and (Z) EF-hand positions on peptides of the high affinity site III of rabbit skeletal troponin C peptide (Marsden et al., 1988).

Table 8. ^1H Resonance Assignments of the modified Parvalbumin peptide sequence at 1°C, pH 6.0.

residue	HN	Hα	Hβ	others
Gly 1		--		
Asp 2	8.44	4.73	2.76	
Ser 3	8.78	4.50	3.91, 3.97	
Asp 4	8.44	4.73	2.76	
Gln 5	8.43	4.30	2.07, 2.18	γH 2.40 δH 7.02, 7.72
Gly 6	8.63	3.96		
Lys 7	8.25	4.36	1.76, 1.82	γH 1.46, 1.52 δH 1.67, 1.68 εH 3.01
Ile 8	8.43	4.18	1.88	γH 1.22, 1.52 γCH_3 0.93 δCH_3 0.88
Gly 9	8.74	4.00		
Val 10	8.17	4.11	2.10	γCH_3 0.929, 0.932
Asp 11	8.64	4.68	2.70	
Glu 12	8.42	4.20	1.83, 1.93	γH 2.10, 2.17
Phe 13	7.92	4.48	2.97, 3.19	ring H 7.26, 7.34, 7.30

Chemical shifts are relative to internal DSS at 2 C.

Figure 19. 2D NMR NOESY of the fingerprint region of the modified parvalbumin sequence.



Marsden showed that substitution of any of the Asp residues in the (X), (Y), (Z) positions decreased La^{+3} affinity. In comparison to his results, the parvalbumin sequence (89-102) and modified parvalbumin sequence had La^{+3} binding affinities of $1x10^{-5}$, and $1x10^{-3}$ respectively. As can be seen the deletion of the Gly and the change of the Asp to Gln in the (Z) position caused a decrease in the La^{+3} binding affinity of the parvalbumin peptide of ~100 fold.

Structure determinations for the modified parvalbumin EF-hand peptide was pursued in two different ways. The first method started with an extended peptide model using 30 NH-α, NH-NH and 7 oxygen to metal ion distance constraints in 2000 rounds of constrained minimization and 100 psec of constrained molecular dynamics. Analysis was made of the total energy over 100 ps range, and it was found that the energy dropped during first 50 ps and during the last 50 ps of dynamics the energy was rather stable. An average structure from the last 50 ps of dynamics was used as an initial model for Mardigras analysis.

In the second method the extended peptide model with 95 NOE and seven metal to oxygen distance constraints were used in 2000 rounds of constrained minimization and 100 psec of constrained molecular dynamics. A structure was chosen with the best total energy, distance and dihedral agreement to be used as an initial structure for the second approach to MARDIGRAS refinement.

In both methods 95 NOE crosspeak intensities were obtained and normalization of the crosspeak intensies was accomplished with the diagonal intensity of an isolated aliphatic proton resonance on the NOESY. These normalized fractional NOE intensities were then used for Mardigras analysis. Dihedral constraints from $J_3Hn\text{-}\alpha$ coupling constants were not available due to line broadening of many proton resonances close to the La^{+3} binding site.

Correlation times of 0.5 and 0.6 ns were calculated from Gd^{+3} relaxation rates from an EF-hand metal binding peptide from rabbit skeletal troponin C (Marsden et al., 1989)(Gariepy et al., 1985). Due to the similarity in MW and structure compactness, it was assumed that the metal bound modified parvalbumin sequence should have a similar correlation time to that of the EF-hand sequence in rabbit skeletal troponin C. Correlation times of 0.5, 1.0, 2.0 and 3.0 nsec were tried in the MARDIGRAS structure refinements. The 0.5 nsec gave distances that were unreasonably short, the 2.0 nsec gave many distances that were too long to observe in a NOESY of a small peptide. The 1 nsec gave reasonable distances consistent with observed NOESY distances of small molecular weight molecules. In both methods a correlation time of 1 ns was used for the MARDIGRAS refinements. This correlation time gave reasonable distances which generated known $\beta\text{-}\beta$ distances of 1.8 Å. In both methods distances calculated from the

Mardigras refinements were used in generating 20 structures from simulated annealing (SA) as described in materials and methods. Analysis of the 20 SA structures showed 10 structures in both methods with chiral violations which were eliminated. Out of the remaining 10 SA structures six were chosen with the best overall distance, dihedral and chiral agreement.

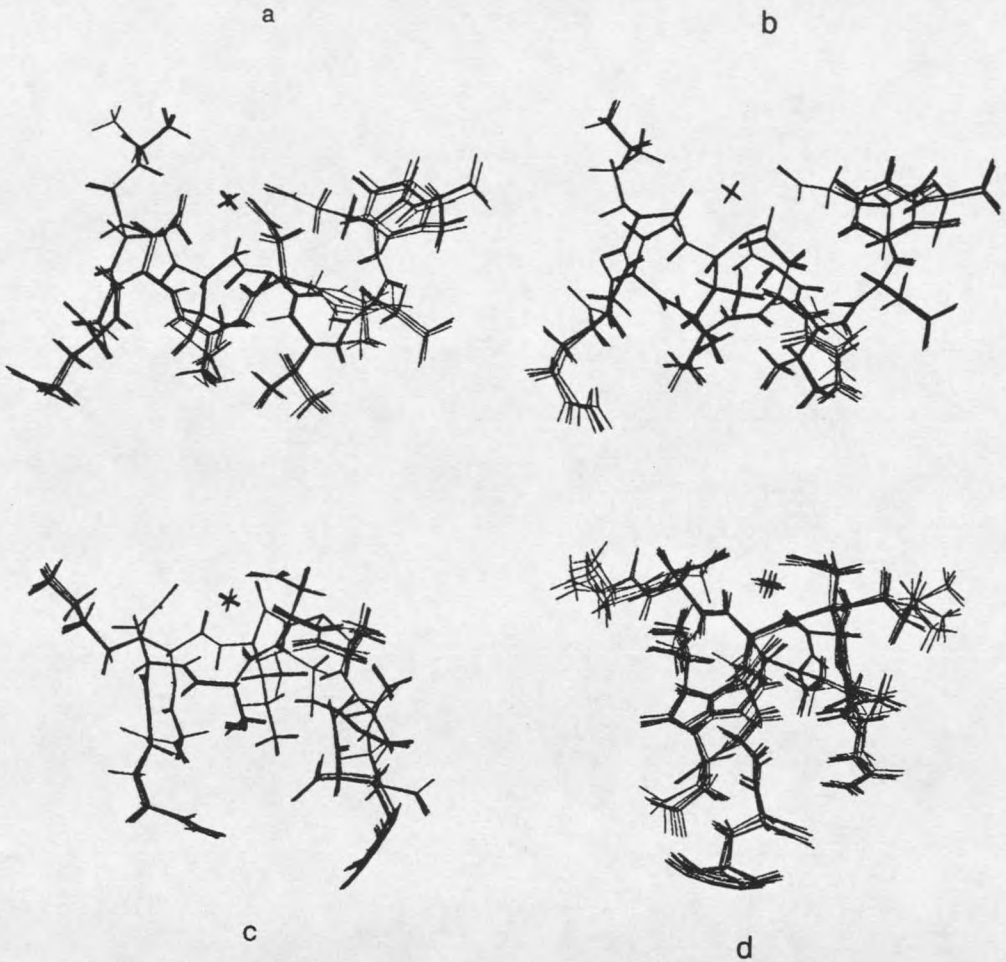
A second round of Mardigras analysis was performed for each of the six structures followed by 100 ps of constrained molecular dynamics. A final iteration of Mardigras was performed for each of the six structures. It was seen that the Q and R values (Withka et al., 1992) did not change and that the calculated distances did not vary between the second and third rounds of Mardigras. The structures are therefore considered converged where further refinement will not improve the structures due to the limited amount of NOESY data. The six sets of the constrained molecular dynamics were then subjected to analysis based on distance, dihedral and chiral agreements. Out of the six sets of structures two had chiral violations, the remaining four sets can be seen in figure 20.

It has been shown that type I and type II reverse turns are present in the EF-hand metal ion binding loops of calmodulin (CaM), parvalbumin (parv) and troponin C (Herzberg et al., 1985). A type I beta turn typically has an Asp or Asn residue in position i, a Pro or Ser residue in

position i+1, an Asp or Asn residue in position i+2, and a Gly residue in position i+3. Type I beta turns have phi and psi dihedrals of the second and third residue in the turn near (-60,-30), and (-90,0). Type III beta turns can have almost any residue in the four positions and have phi psi dihedral angles in the second and third positions near (-60,-30), and (-60,-30) (Wilmot et al., 1988)(Chou et al., 1977). In the 12 residue EF-hand loops of CaM, parv, and TnC the type I beta turn is found from residue one to four and the type III beta turn is found from residue nine to twelve. In most of these classical EF-hand loops Asp and Asn/Asp residues in the first and third positions which are typical for a beta I turn. For the beta-III turn there are a greater variety of residues found in the four positions.

The residue positions in the modified parvalbumin that are expected to have a potential beta I turn in standard EF-hand are from 2Asp to 5Gln. The modified parvalbumin peptide has an Asp in position i, a Ser in position i+1, and an Asp in position i+2 which are residues that have a high probability of forming a beta I turn. The phi psi dihedral angles found in all of the best fitting structures for position i+1, and i+2 in the four structure sets of modified parvalbumin can be seen in table 9.

Figure 20. Four best sets of modified parvalbumin structures (a-d); after SA, MD and Mardigras.



94

Table 9. (phi-psi) dihedral angles of the four sets of structures for residues 3Ser, and 4Asp.

Turn	$(phi\text{-}psi)_2$	$(phi\text{-}psi)_2$
beta I	-60, -30	-90, 0
structure a	-86, 99	-117, 88
structure b	-117, -66	42, 38
structure c	-136, 94	-104, 72
structure d	-144, 101	-128, 71

The phi-psi values for the second and third residues in the potential β bend in residues 2-5, do not match well with the phi-psi values in a typical beta-I turn. The $d_{N-N}(i,i+2)$ distance for a beta-I turn is 3.8 Å, and in the modified parvalbumin peptide it ranges from 5-6 Å. In the modified parvalbumin the sequence from Gly-9 to Glu-12 might be expected to be a potential beta-III turn. In a beta-III turn a specified starting residue position is not as important as in a beta-I turn, but there is a phi-psi dihedral angle range that typifies a beta-III turn which can be seen in table 10.

Structures a-d do not resemble the typical phi psi dihedral angles of a beta-III turn. However residues 9Gly to 12Glu in the modified parvalbumin peptide do form a turn in order for the sidechain carboxyl oxygens of 12Glu to be in a position to bind to the metal ion.

Table 10. Phi-psi dihedral angles of 10-Val, and 11-Asp in the four structure sets of modified parvalbumin peptide.

Turn	$(phi\text{-}psi)_2$	$(phi\text{-}psi)_3$
beta-III	-60, -30	-60, -30
structure a	-109, 69	51, 108
structure b	-118, 70	37, 120
structure c	58, 127	16, 77
structure d	-60, -40	61, 90

A possible reason for the deviation from a beta-I turn to a more extended turn in residues 2-5 in the modified parvalbumin peptide, may be due to the deletion of Gly (Z-1) and substitution of Gln for the Asp (Z). Structures from constrained molecular dynamics using all 95 NOE distances without the seven metal to oxygen distances showed that extended structures were possible that could agree reasonably with the NOE distance restraints (results not shown). This implies that the seven metal to oxygen distance constraints were an important driving force in the structure of the La^{+3} complex of the modified parvalbumin.

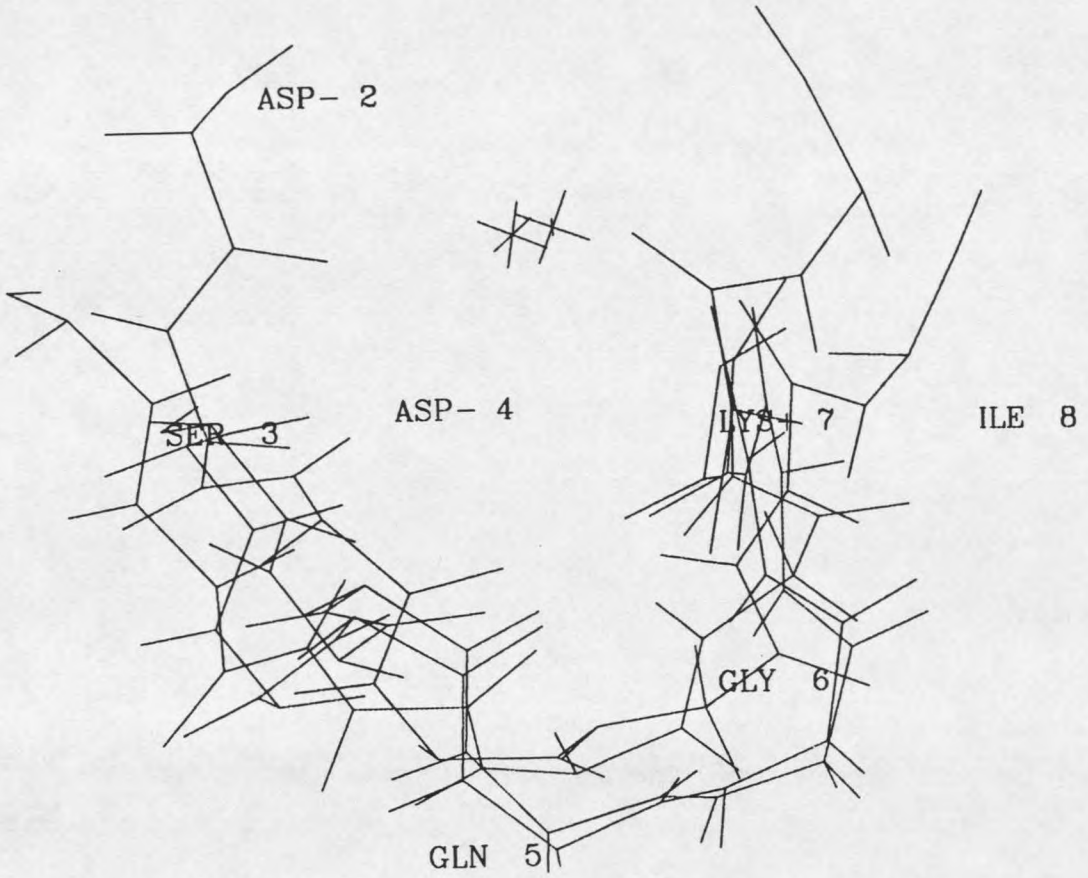
Structures of the region from 2Asp to 8Ile from all the four sets in figure 20 were found to agree rather well when they were compared to find the best potential structure for attempted docking onto BR. Figure 21 shows the superposition of the best structures from the four sets of structures in figure 20 as carried out with the superimpose utility in InsightII (InsightII User Guide, 1993). All

structures showed a similar broad turn from 2Asp to 8Ile and superimposed rather well. The utility employed does not provide a RMSD for the superposition.

The 2Asp to 8Ile region of the modified parvalbumin peptide was used to dock onto BR in position 102-108 between the C and D helices. The positions of the helices are fairly well known from the electron microscopy structure of Henderson and et al (1990). The helices were fixed in space, and only the loops between helices C&D and E&F were allowed freedom to move during dynamics. The La^{+3} complex of 2Asp-8Ile fit very well on the fixed C and D helices between 102-108 with no distortion or adjustment needed

The segment of the modified parvalbumin peptide docked onto position 102-108 of BR has five oxygens that bind to the metal ion, two each from 2Asp and 4Asp and one from the backbone oxygen of 7Lys. An EF-hand binding loop normally has a seven coordination to a metal ion. As previously explained there is evidence that a residue with a sidechain carboxyl that is distant in the sequence may be able to fill the (-Z) EF-hand binding position (Vyas et al., 1987). 161Glu and 166Glu are potential residues to complete the EF-hand coordination located in the loop between helix E and F in the tertiary structure of BR. The 161Glu or 166Glu may be close enough to contribute to bind a metal ion on the C-D loop of BR. After the La^{+3} complex of 2Asp to 8Ile was docked onto the BR helices as discussed, constrained

Figure 21. Superposition of the best structures from figure 20, showing the backbone atoms of residues 2Asp-8Ile.



molecular dynamics and energy minimization was performed using 41 NOE distance constraints located between the 2Asp to 8Ile residues including five oxygen to metal ion constraints. Two constraints from the 161Glu or 166Glu sidechain carboxyl oxygens to the metal ion were also added. Results showed that the sidechain carboxyl oxygens of 161Glu could come within 2.6 Å of the metal ion in the C-D loop. The sidechain carboxyl oxygens from 4Asp fit slightly above the plane of 4 oxygens from 102Asp, 107Lys, and 161Glu, and one oxygen from 161Glu slightly below the plane (figure 22).

Constrained molecular dynamics was performed again on the BR system but two constraints from the sidechain carboxyl oxygens of 166Glu to the metal ion were substituted for the 161Glu. Results showed that the sidechain carboxyl oxygens of 166Glu could come within 2.7 Å of the metal ion in the C-D loop. The sidechain carboxyl oxygens from 166Glu are slightly above a plane of the four oxygens from 102Asp and 104Asp, with the oxygen from 107Lys below the plane (figure 23). This shows a very crowded metal ion binding environment on one side with poor coordinates of the metal ion on the other side.

In both the 161Glu and 166Glu cases there are no distance violations greater than .34 Å for the 41 NOE distance constraints along the 102Asp to 108Ile sequence. In both cases there is no dihedral or distance evidence of a specific beta turn, but an extended turn in this region is

evident. While the metal ion binding environment is most likely not acceptable for the 166Glu case, the 161Glu case has a very plausible metal ion coordinate. In the 161Glu coordinate the metal ion binding environment is distorted from a normal pentagonal bipyramidal binding geometry of an EF-hand system and the geometry of the metal site is a distorted pentagonal bipyramid. It seems plausible that the sidechain carboxyl oxygens from 161Glu contribute to binding a metal ion located in the loop between helix C and D in BR.

Our proposal for a metal ion binding site on the loop between helix C and D in BR with completion of metal coordination by 161Glu on loop EF can be used as a model for further work. There is currently little evidence on the structure of the surface loops of BR and more sophisticated solid state NMR experiments could be used to test this model. For example Rotational Resonance (RR) or Rotational Echo Double-Resonance (REDOR), with the aid of specific isotopic labelling (S.O. Smith et al., 1992) could be carried out to measure distances between residues in the CD loop and between the CD and EF loops with and without metal ions.

Figure 22. Potential metal coordination site in BR, with a distant sidechain carboxyl from 161Glu located in the loop between helices C-D, along with sidechain carboxyl oxygens from the 102Asp, 104Asp and the backbone carbonyl oxygen of 107Lys residues.

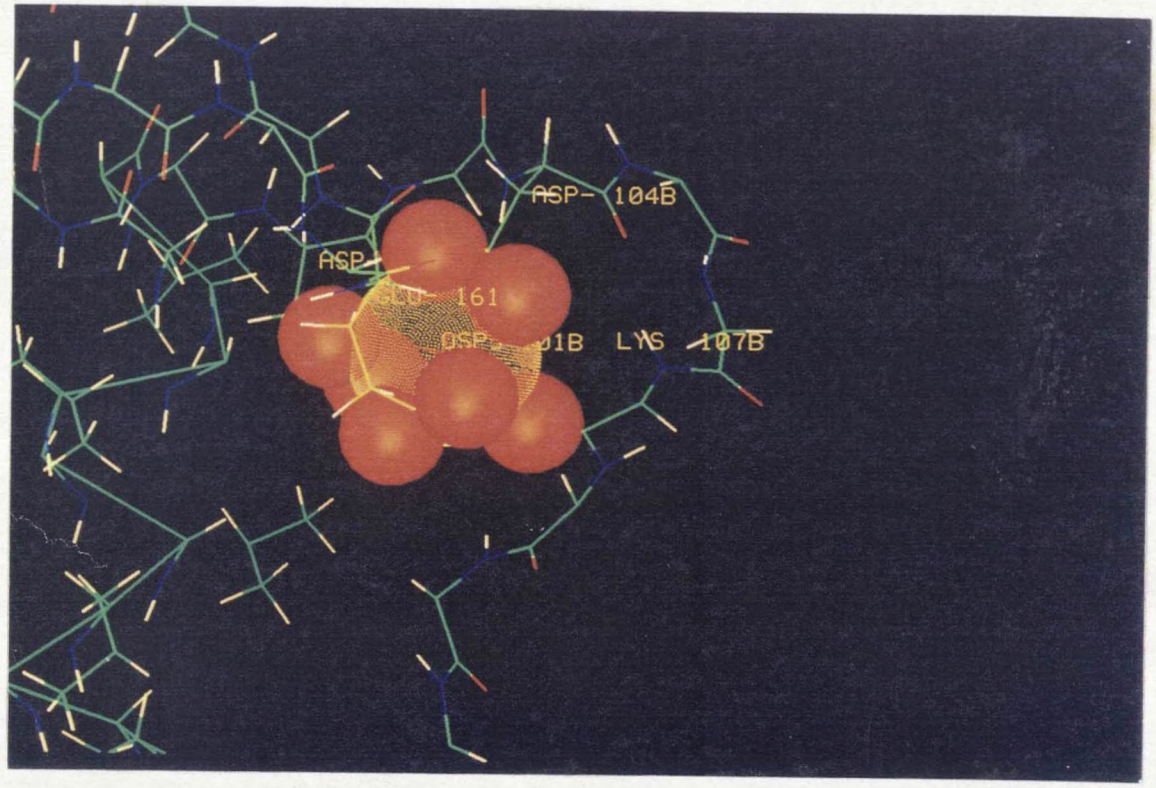
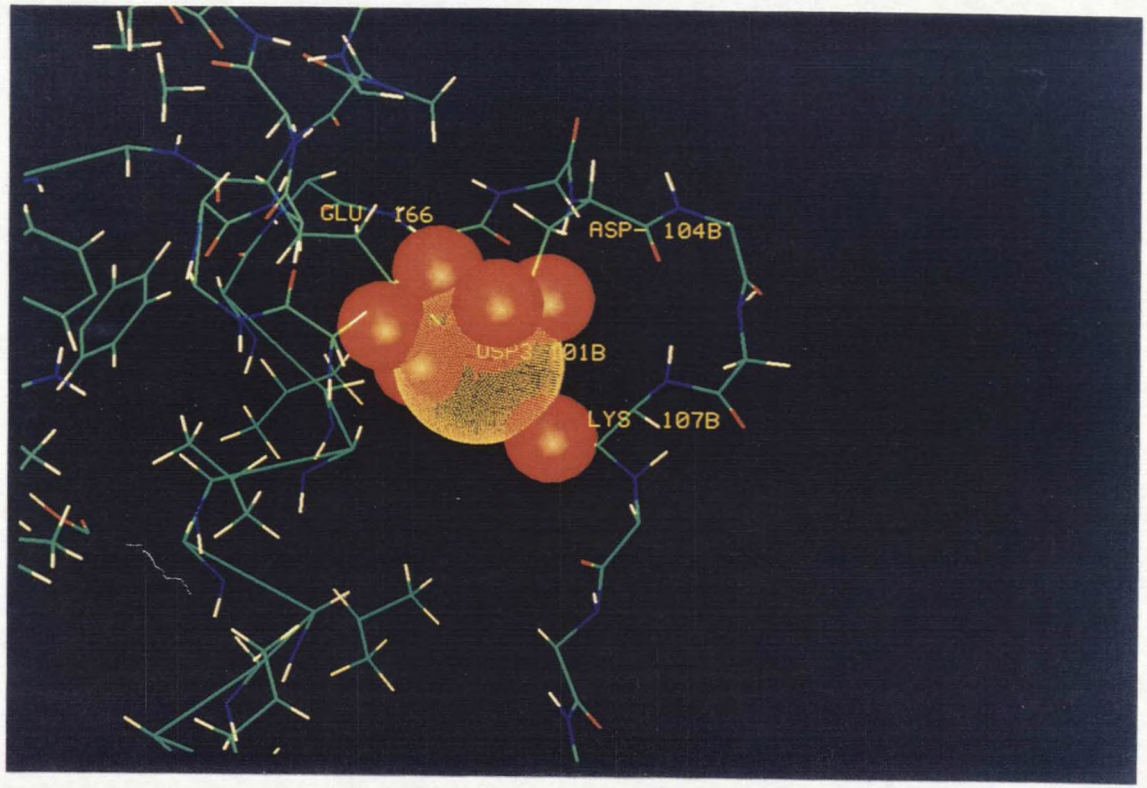


Figure 23. Potential metal coordination site in BR, with a distant sidechain carboxyl from 166Glu located in the loop between helices C-D, along with sidechain carboxyl oxygens from the 102Asp, 104Asp and the backbone carbonyl oxygen of 107Lys residues.



Chapter 4

STRUCTURE OF PROTEIN-PROTEIN CONTACT REGIONS USING TRANSFERRED-NOESY

Protein-protein and interactions are at the heart of many different mechanisms in biology. There are very few cases where protein-protein complexes have been crystallized and studied by x-ray crystallography. NMR has potential as a tool for studying protein-protein interfaces using peptides from one of the proteins that block the protein-protein interactions (Dratz et al., 1993). NMR has been used to study interactions such as enzyme-substrate, hormone-receptors, and antibody-antigen. The large size of macromolecules or receptors results in slow tumbling in solution on the NMR time scale. 2D NMR can provide little if any direct structural information on macromolecules larger than 8,000-10,000 daltons. With the advent of 3D and 4D NMR techniques in the late 80's and early 90's structures of molecules up to 30,000 daltons can be studied with NMR as long as complete substitution of all ^{15}N and ^{13}C sites can be accomplished (Clore et al., 1991). In most cases biological receptors of interest are larger than the upper limit for full NMR structure determination. In addition biological receptors are often membrane bound with very slow molecular

motions which broaden NMR lines to the point that liquid NMR is not useful for structural determination. Transferred NOESY (TR-NOESY) allows the study of the conformation of many different types of small molecules binding to large proteins (James, 1976) (Campbell et al., 1993) including peptides bound to large membrane bound receptors (Dratz et al., 1993).

TRNOE Theory

The continuous wave TRNOE experiment was originally proposed by Balaram, Bothner-By and coworkers (Balaram et al., 1972)(Balaram et al., 1972) and was developed by Clore and Gronenborn (Clore et al., 1983)(Clore et al., 1982). TRNOE was developed to study the conformations of ligands bound to binding proteins. The TRNOE makes use of chemical exchange between nuclei in the bound and free states. Information on the bound state ligand conformation is established via proton-proton cross relaxation and this information is transferred to the sharp NMR spectrum of the free ligand by chemical exchange. TRNOE obtains the information on conformation of a ligand bound to a receptor system by observing negative NOE's on the free ligand.

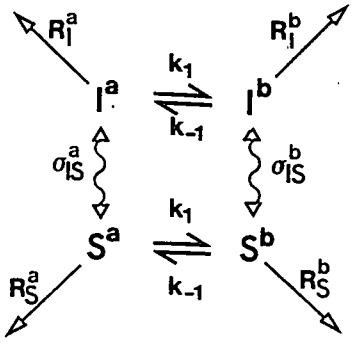
In order to better understand the applications and limitations of TRNOE in determining the conformation of ligands in chemical exchange with large protein systems, the theory of the chemical exchange and nuclear relaxation model will be briefly presented below.

The kinetic reaction scheme that will be used to model ligand-protein binding is:

$$a \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} b$$

Where exchange is considered between two environments a and b with forward and backward rate constants k_1 and k_{-1}. Figure 24 shows the nuclear spin species in a two spin system, I and S, with exchange between free ligand (a) and protein bound ligand (b).

Figure 24. Species and rate constants in a two spin system two site undergoing exchange (Neuhaus and Williamson, 1989).



where I^a, I^b are spin I in a and b environments, and S^a, S^b are spin S in a and b environments. R_I^a, R_I^b, R_S^a, and R_S^b are the relaxation constants for spins I and S in the a and b environments, and σ_{IS}^a, σ_{IS}^b are the cross relaxation rates between spins I and S in a and b environments. The forward and reverse exchange rates are again k_1, and k_{-1}.

Modified McConnell equations are used in order to find

the values of NOE enhancement for I^a, and I^b upon irradiation of spin S^a, and S^b. In the steady state case, the relation is (Neuhaus and Williamson, 1989).

$$f_I^a(S) = \frac{\sigma_{IS}^a(R_I^b + k_{-1}) + k_1\sigma_{IS}^b}{(R_I^a + k_1)(R_I^b + k_{-1}) - k_1k_{-1}} \qquad 27$$

$$f_I^b(S) = \frac{\sigma_{IS}^b(R_I^a + k_1) + k_{-1}\sigma_{IS}^a}{(R_I^a + k_1)(R_I^b + k_{-1}) - k_1k_{-1}} \qquad 28$$

In slow chemical exchange, when (k<<R) on the T_1 timescale, equations 27-28 reduce to:

$$f_I^a(S) = \frac{\sigma_{IS}^a}{R_I^a} \qquad 29$$

$$f_I^b(S) = \frac{\sigma_{IS}^b}{R_I^b} \qquad 30$$

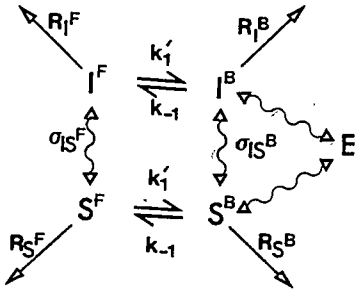
Equations 29-30 indicate that if chemical exchange is slow on the T_1 timescale, the average NOE enhancement for spin I occurs separately for the two environments a and b. In fast chemical exchange when (k>>R) on the T_1 timescale there are average resonances for spin I in the a and b environments and equation 27-28 reduce to:

$$f_I^a(S) = f_I^b(S) = \frac{k_{-1}\sigma_{IS}^a + k_1\sigma_{IS}^b}{k_{-1}R_I^a + k_1R_I^b} \qquad 31$$

Equation 31 indicates that in the fast exchange case, the average NOE enhancement occurs as a weighted average of the two environments a and b.

The two spin system with exchange can be expanded to consider cross relaxation (or spin diffusion or spin polarization leakage) to the macromolecular receptor (E).

Figure 25. Two spin system involving a ligand molecule undergoing exchange (free and bound states) with a receptor molecule E (Neuhaus and Williamson, 1989).



Campbell and Sykes represent leakage of spin polarization to E with a leakage parameter (Campbell et al., 1991). Other treatments such as Zheng and Post (1993) and Nirmala et al (1992) treat specific transfers from the ligand to the macromolecular receptor.

The time evolution of the longitudinal magnetizations, of the ligand exchanging with a macromolecular binding site assuming that only the ligand resonances are observed, can

be represented as

$$\frac{d\mathbf{m}}{dt} = -\mathbf{R}\mathbf{m} \qquad 32$$

$$\mathbf{m} = \begin{bmatrix} m_{IB} \\ m_{SB} \\ m_{IF} \\ m_{SF} \end{bmatrix} \qquad \mathbf{R} = \begin{bmatrix} R_{II}^{B}+k & R_{IS}^{B} & -kf & 0 \\ R_{SI}^{B} & R_{SS}^{B}+k & 0 & -kf \\ -k & 0 & R_{II}^{F}+kf & R_{IS}^{F} \\ 0 & -k & R_{SI}^{F} & R_{SS}^{F}+kf \end{bmatrix}$$

$$f = \frac{[EL]}{[L]}$$

Where **m** represents the deviations of magnetization from equilibrium the **R** matrix contains relaxation and exchange rates of spins I and S in the bound and free states, and f is the ratio of bound ligand to free ligand in equilibrium. The off rate k_{-1} is shown as k for simplicity of notation (Lippens et al, 1992).

The **R** matrix can be separated into exchange **E** and relaxation **W** matrices.

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_b & 0 \\ 0 & \mathbf{W}_f \end{bmatrix} \qquad \mathbf{E} = \begin{bmatrix} k_1'\mathbf{1} & -k_{-1}\mathbf{1} \\ -k_1'\mathbf{1} & k_{-1}\mathbf{1} \end{bmatrix} \qquad 33$$

where $k'_1 = k_{-1}f$ and **1** and **0** are the unit and null matrices respectively. The solution of the above equations can be shown to be

$$\mathbf{m}(t) = \exp(-\mathbf{R}t)\mathbf{m}(0) \qquad 34$$

Equation 32 can be expanded through several steps into

$$\frac{dm}{dt} = \begin{bmatrix} \mathbf{m}_b + \mathbf{m}_f \\ p_b\mathbf{m}_f - p_f\mathbf{m}_b \end{bmatrix} = -\left[\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (k_1 + k_{-1})\mathbf{1} \end{bmatrix} + \begin{bmatrix} p_b\mathbf{W}_b & \mathbf{W}_f - \mathbf{W}_b \\ p_b p_f(\mathbf{W}_f - \mathbf{W}_b) & p_b\mathbf{W}_f + p_f\mathbf{W}_b \end{bmatrix}\right] X \begin{bmatrix} \mathbf{m}_b + \mathbf{m}_f \\ p_b\mathbf{m}_f - p_f\mathbf{m}_b \end{bmatrix} \qquad 35$$

where p_b and p_f are fractional concentrations of the free and bound species. Due to the presence of exchange the resulting matrix is unsymmetric and is not easily solved. This is the case for intermediate chemical exchange. However, if the system is in fast exchange, then the term $(k_1 + k_{-1})\mathbf{1}$ in the above equation will dominate because the system will reach equilibrium quickly. Therefore the bottom terms, which describe the departure of the system from equilibrium, will have little effect on the system and can be ignored. Therefore equation 35 becomes

$$d\frac{\mathbf{m}}{dt} = [p_b\mathbf{W}_b + p_f\mathbf{W}_f \;\; \mathbf{W}_f - \mathbf{W}_b]\, X \begin{bmatrix} \mathbf{m}_b + \mathbf{m}_f \\ p_b\mathbf{m}_f - p_f\mathbf{m}_b \end{bmatrix} \qquad 36$$

where the deviation from equilibrium is essentially zero as seen below

$$p_b\mathbf{m}_f - p_f\mathbf{m}_b = 0 \qquad 37$$

and the resulting equation will be as such

$$\frac{dm}{dt} = (p_b W_b + p_f W_f)(m_b + m_f) \qquad 38$$

Equation 38 shows that in the fast exchange limit, the magnetization decays by the concentration weighted sum of the bound and free relaxation matrices (Landy and Rao, 1989). So in fast exchange where the NMR spectra of both species are likely indistinguishable, the sum of the free and bound relaxation matrices can be measured by the time evolution of the magnetization of the system.

The fundamental basis of the TRNOE is that the cross relaxation of the receptor bound ligand system (σ^b) in the bound state is much larger than the cross relaxation of the ligand of the free state (σ^f). In the case of fast exchange, σ^b will dominate the overall cross relaxation $\langle\sigma_{IS}\rangle$ in the ligand-receptor system. Since the overall cross relaxation of the system in fast exchange is the weighted average of the bound and free species it can be represented as follows:

$$\langle\sigma_{IS}\rangle = p_f\sigma^f_{IS} + p_b\sigma^b_{IS} \qquad 39$$

where p_f, and p_b are the fractions of the free and bound species respectively. If the bound cross relaxation dominates in the system then

$$|p_b\sigma^b_{IS}| >> |p_f\sigma^f_{IS}| \qquad 40$$

In order to fulfill this inequality the influence of

parameters such as affinity for the formation of the receptor-ligand complex, tumbling rates of the free and bound species, and rate of exchange between the bound and free ligand must be addressed. However, in fast exchange the contribution of the free peptide signal can be subtracted from the observed cross relaxation rates to obtain the bound cross relaxation rates. The conditions needed for successfully applying TRNOE to a ligand receptor system will be discussed in the following TRNOESY section.

TRNOESY

TRNOESY is approximately equivalent to carrying out a large number of TRNOE experiments with evenly spaced irradiation frequencies. Even though TRNOESY has inherently somewhat lower sensitivity than TRNOE, the baselines tend to be better in TRNOESY and excitation is much more selective. Therefore, better data is routinely obtained with TRNOESY. In the past 10 years TRNOESY has been used to study the binding between hormones and receptors (Lippens et al., 1992); enzymes and substrates (Ni et al., 1990)(Williams et al., 1991)(Meyer et al., 1988); peptides and phospholipid bilayer (Milon et al., 1990)(Wang et al., 1993); antigens and antibodies (Cung et al., 1991)(Scherf et al., 1992)(Glaudemans et al., 1990); and peptides and receptors (Dratz et al., 1993).

Chemical exchange rate, mixing time, and concentration of free and bound ligand have effects on the TRNOESY. Early

studies by Clore and Gronenborn indicated three regions of kinetic exchange. They found the fast kinetic exchange region occurred when the off rate of the ligand was at least ten times faster than the spin lattice relaxation the magnitude of the TRNOE on two protons of the ligand is directly proportional to the bound cross relaxation rate. Later analysis has shown that the off rate of the ligand has to be faster than this and must be at least ten times faster than the fastest bound cross relaxation rate in order for the simplifications of the fast exchange to be used (Lippens et al., 1992). The intermediate kinetic exchange region occurred if the off rate was just faster than the spin lattice relaxation. The TRNOE in the intermediate exchange case was still approximately proportional to the bound cross relaxation rate if the ligand was present in large excess. In the case of slow exchange between the bound and free ligand relative to spin lattice relaxation the TRNOE was very small.

One of the first questions that can be addressed is what type of experimental information can be gathered from TRNOESY data. Anglister categorized the types of TRNOE peaks that occur in an antibody-peptide system: a.) chemical exchange between the free and bound peptide resonances. b.) chemical exchange between free and bound protein resonances. c.) magnetization transfer within the protein. d.) intermolecular magnetization transfer between

protein and free peptide via the bound peptide. e.) intramolecular magnetization transfer within the bound peptide via exchange with the free peptide. Category a and b will only be seen in slow exchange cases where the free and bound spectra are distinct. In the case of magnetization transfer within the protein, (protein-protein NOE's), resonances will be broad and hard to detect. Only categories d and e constitute true TRNOE crosspeaks. Category d has broad peaks in the protein dimension. Category e is the case which gives information on the bound conformation of peptides, while category d could be useful for determining interactions between a peptide and protein system if the protein is of modest molecular weight and peaks of the protein are not too broad (Campbell et al., 1993).

As explained above there are two true categories of TRNOESY crosspeaks (category d and e above). It is useful to be able to differentiate intermolecular and intramolecular TRNOESY crosspeaks. It has been shown that TRNOESY intensity is dependent on the ligand/protein ratio (Campbell et al., 1993). This dependency of the TRNOESY intensity on the ligand/protein ratio can be used to differentiate intramolecular and intermolecular TRNOESY crosspeaks. As the protein/ligand ratio increases at constant protein concentration the intramolecular fractional TRNOESY intensity increases (Ni et al., 1990) while

intermolecular fractional TRNOESY intensity stays constant (James, 1976) (Anglister et al., 1989).

In fast exchange as the protein/ligand ratio increases the intramolecular TRNOESY intensity increases (Ni et al., 1990)(James, 1976)(Anglister et al., 1989). This can be seen using $p_f + p_b = 1$ and rearranging equation 39 to:

$$\sigma_{IJ}^{obs} = \sigma_{IJ}^{f} + p_b(\sigma_{IJ}^{b} - \sigma_{IJ}^{f}) \qquad 41$$

where the only independent variable is p_b.

The reaction system that will be used for protein ligand binding in the following discussion is shown:

$$L + P \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} LP \qquad 42$$

$$K_A = \frac{[LP]}{[L]\,[P]} \qquad 43$$

where L is the ligand, P the protein, LP the ligand protein complex, and K_A is the association constant. The intermolecular TRNOESY depends on the total concentration of ligand-protein complex [LP]. Under typical NMR conditions the concentration of ligand is in excess of the dissociation constant $K_D=1/K_A$ so virtually all the protein will be in the LP complex. The observed intramolecular TRNOESY increases with p_b, but p_b decreases with increasing ligand

concentration. The intermolecular and intramolecular TRNOESY can be distinguished by keeping the protein concentration constant and varying the ligand concentration. The intermolecular TRNOESY will stay constant (as long as virtually all of the protein is in LP) or go down with decreasing ligand whereas the intramolecular TRNOESY (and p_b) will increase as the ligand concentration decreases.

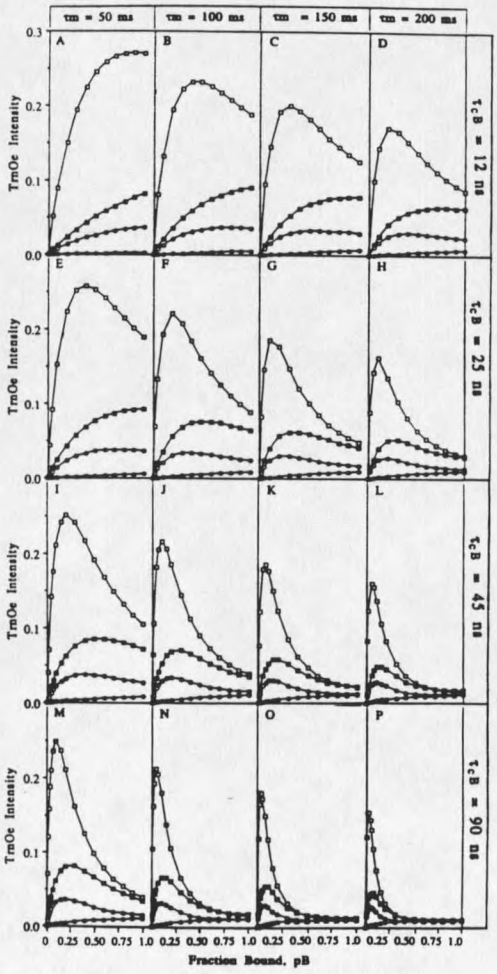
Once it has been established that a TRNOESY interaction has occurred between a ligand and a receptor system, the question of specific binding interaction between the ligand and the receptor arises. Gierasch showed that TRNOESY interactions occurred in a farnesyltransferase-hexapeptide system. TRNOESY experiments were then done in which the sequence of the hexapeptide system was scrambled which still resulted in some weak TRNOESY crosspeaks (Gierasch et al., 1993). Landy and Rao showed that the interaction of MgADP and creatine kinase give TRNOESY interactions. In the MgADP-creatine kinase system it was shown that upon addition of MgADP of up to 6mM the percent NOE was independent of MgADP concentration, but from 6 mM to 10 mM the percent NOE increased which must be due to nonspecific effects. Experiments were performed between MgADP and other proteins which should show no specific interactions to MgADP. Weak TRNOE interactions were in fact seen between MgADP and the nonspecific proteins. This experimental evidence indicates that nonspecific binding of the ligand to proteins can occur

and must be taken into account by using biological activity controls such as; making measurements over a range of ligand to protein concentrations (Murali et al., 1993) using peptides that bind tightly and inhibit the TRNOESY; and/or using inactive peptides with slightly modified sequences to observe nonspecific effects on the TRNOESY. For example, excellent biological acitivity controls were available to show specificity on a GTP binding protein peptide interacting with rhodopsin (Dratz et al., 1993). For example, in this case amidinating the terminal carboxyl end of the peptide led to the loss of all biological activity and all TRNOESY signals.

Campbell and Sykes have performed TRNOESY simulations of a 12 residue peptide from troponin I (TnI) in chemical exchange with the muscle protein troponin C (TnC) some of which are reproduced in figure 26 (Campbell et al., 1993). The relationship of TRNOESY buildup to bound correlation times, concentration of bound ligand, and mixing times are illustrated in these calculations. It can be seen in the series (A-D) in figure 26 that as mixing time (τ_m) is increased the TRNOESY buildup rate deviates from linearity as a function of fraction bound at a smaller fraction bound (p_b). TRNOESY builds up faster at smaller p_b since the cross relaxation rate between protons in the bound state is faster than the cross relaxation in the free state, ($W_b >> W_f$). It is also seen that as the bound correlation

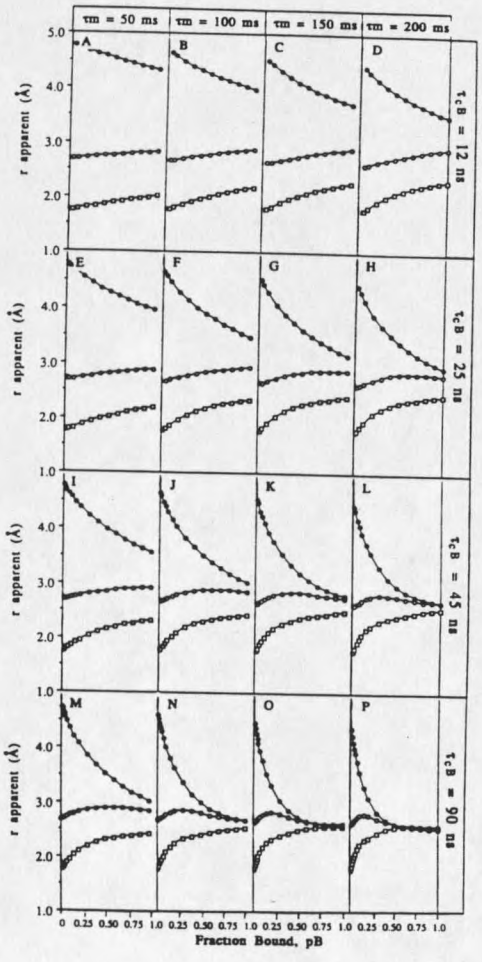
time (τ_b) increases in a series (D-P), that the TRNOESY intensity builds up faster at smaller p_b and decays faster at higher p_b. In cases both of large τ_m and p_b the TRNOESY magnetizations approach the same intensities which indicates extensive spin diffusion. Our experiments are closest to panel H in figure 26 at the low fraction bound (0.01-0.10) where the NOESY intensities are quite linear.

Figure 26. TRNOE intensity buildup in relation to concentration of bound ligand, mixing time, and correlations time (Campbell and Sykes, 1993) for protons at different distances.



The trend for increasing spin diffusion with increased fraction bound can be seen in figure 27, where the TRNOESY intensities of figure 26 are converted into distances. It can be seen that at higher τ_m and τ_b that the distances converge to approximately 2.5 Å. Both figures indicate that TRNOESY analysis should be done at short τ_m and small p_b.

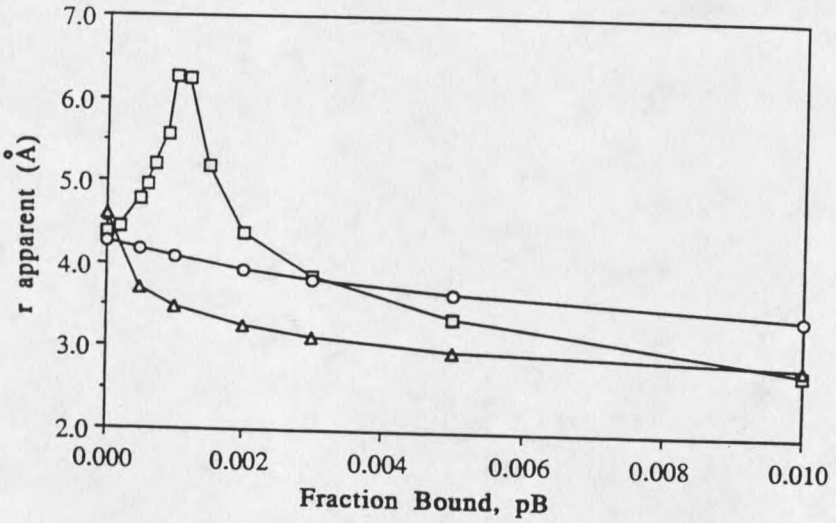
Figure 27. TRNOE intensities from figure 26 converted to distance in relation to fraction of bound ligand, mixing time, and correlation time (Campbell et al., 1993.



A source of potential concern for the analysis of the

TRONESY is how the free conformation of the peptide contributes to the bound data. Figure 28 shows a simulation of apparent distance (r) versus fraction peptide bound (uncorrected for the free peptide) of an arginine proton pair where the correlation times of the free peptide were picked to be 0.2, 0.4, and 0.8 ns. The distance of this proton pair in the free is 4.41 Å. At a p_b of ~ .01 the r apparent is nearing the bound distance of 2.46 Å.

Figure 28. Simulation of apparent distance (r) for R112NH-R113NH proton pair versus fraction bound peptide (p_b) for correlation times of: □, τ_{cf}=0.2 ns; △, τ_c^f=0.4 ns; O, τ_c^f=0.8 ns (Campbell and Sykes., 1993) with no correction for the contribution of free peptide.



The curve for the τ_c of 0.2 ns approaches the bound distance the fastest. The discontinuity of the 0.2 ns curve is due to the NOE of the free peptide passing from the positive to the negative region. Therefore it is seen that the TRNOE

distance is dominated by the bound species even at modest p_b (> 0.003) (Campbell et al., 1993). Furthermore, where it can be shown that the peptide is in fast exchange on the cross relaxation time scale it is possible to correct for any contribution of the free peptide NOESY as will be further discussed later. Correction for the contribution of the free peptide makes the distance calculation such as shown in figure 28 much more constant.

Protons on the protein may affect the apparent distances of the bound ligand. Due to efficient cross relaxation between protein protons, it would be expected that indirect effects of protein protons could be treated as a single relaxation sink as shown by Campbell and Sykes (1993) and would have an overall small effect but such may not always be appropriate. Simulations were done on two spins exchanging with a complex with a third spin using a full matrix relaxation analysis called by the authors 2D Exchange Transferred NOESY (ET-NOESY) but which is equivalent to TRNOESY. The rate matrices simulated proton spins from the ligand and protein. The simulations showed that the protein protons in certain geometries can have an effect on the bound ligand TRNOE intensities (Zheng and Post, 1993).

The effect of cross relaxation (spin diffusion) from protein protons on ligand TRNOESY intensities cannot typically be compensated by NOE distance refinements

protocols such as MARDIGRAS (Borgias et al., 1990) since the positions of receptor protons in space are normally not known. Zheng and Post conducted experiments on the NADH lactate dehydrogenase system and showed one NADH distance was affected by a protein proton. This effect could be established since there was a crystal structure for the NADH lactate dehydrogenase complex and therefore a complete relaxation matrix analysis could be carried out. Indirect effects from the protein protons was for the most part negligible. The possibility of an indirect effect of a nearby protein proton depending on its geometry in space in relation to the protons on the ligand must be recognized. If certain distances in the bound structure are inconsistent with the others these will be rejected in Mardigras refinement.

Nirmala has done TRNOESY simulations on a four spin system in which I and S are spins on the ligand and X and Y are spins on the protein in which Y is a faster relaxing proton sink, whereas X sites are approximately the same all around the protein. Simulations showed that high peptide/protein ratio minimized the influences of protein relaxation as previously pointed out by Campbell and Sykes (1991). The larger number of relaxation sinks near the binding site of the ligand in the protein resulted in a decrease in NOE intensity by a relaxation pathway of IB→XB→YB. These simulations also showed that more Y

relaxation sinks moved the TRNOE maximum towards shorter mixing times. This would indicate that the maximum TRNOE intensity of the I and S spins of the bound ligand is dependent on the close contacts from X and Y spins on the protein (Nirmala et al., 1992). Experiments with high peptide to protein ratios and mixing times less than about 200 msec are predicted to be relatively free of protein relaxation effects (Nirmala et al., 1992).

Besides the effect of protein protons on the TRNOESY of the bound ligand, the longitudinal relaxation (ρ_f) of ligand protons in the free ligand can affect the TRNOESY of the bound ligand. The cross relaxation (σ_f) of the free peptide may be near the extreme narrowing limit ($\omega\tau_c=1$) where $\sigma_f \sim 0$, but dipolar T_1 relaxation, ρ_f, on the ligand is a very significant effect on the NOESY intensities. The effect of T_1 from the ligand can be easily compensated, by dividing the TRNOESY crosspeak intensities by the diagonal intensity of a resolved proton in the NOESY.

As shown by Clore and Gronenborn (Clore et al., 1982), TRNOE distance determination from a system in slow chemical exchange is not reliable and will tend to give erroneous results. A slow exchange system is usually due to a slow off rate (k_{off}). As a protein becomes bigger or moves slower, cross relaxation rates increase. One way to increase the off rate of a system is increase the temperature or increase the salt concentration in the system

in which electrostatic interactions between the protein and ligand may help to increase the k_{off}. The disadvantage in manipulating the k_{off} rate in this manner is that the ligand could form secondary nonspecific binding sites to the protein (Lippens et al., 1992). In the case of peptide ligand the peptides can be shortened to weaken binding and increase k_{off}, if needed, to obtain fast exchange.

gp91-*phox* carboxyl terminal peptide inhibition of superoxide

Stimulation of human neutrophils with chemoattractants, opsonized particles, and other agonists results in the release of superoxide anion (O_2^-) into phagolysosomes or into the extracellular environment (Badwey et al., 1980)(Baggiolini et al., 1993). The initiation of O_2^- production results from the activation of an inducible plasma membrane-bound NADPH oxidase, which is an electron transport system that transfers high energy reducing equivalents from NADPH to O_2 to form O_2^- (Babior, 1987)(Clark, 1990)(Morel et al., 1991). The activation of this system involves the assembly of membrane-bound and cytosolic components to form the active complex in the plasma membrane (Babior, 1987)(Clark, 1990)(Morel et al., 1991). The plasma membrane-associated component directly mediating electron flow is a low-potential b-type cytochrome (a.k.a. cytochrome b_{558}, cytochrome b_{559}, or cytochrome b_{-245}) (Segal, 1989). In neutrophil membranes, a low molecular weight GTP-binding protein, Rap1A, is also associated with

cytochrome b (Quinn et al., 1989) although its function in the NADPH oxidase is currently unknown. The cytosolic components required for NADPH oxidase activity include p47-*phox*, p67-*phox*, and a second low molecular weight GTP-binding protein, Rac (Volpp et al., 1988)(Nunoi et al., 1988)(Knaus et al., 1991)(Abo et al., 1991). Upon activation of the NADPH oxidase, all three of these cytosolic proteins translocate to the plasma membrane where they associate with the membrane-bound components to form the active O_2^--generating complex (Clark et al., 1990)(Quinn et al., 1993).

Cytochrome b is a heterodimer of a glycosylated 91 kDa subunit (gp91-*phox*) noncovalently bound to a nonglycosylated 22 kDa subunit (p22-*phox*) (Parkos et al., 1987)(Parkos et al., 1988) and appears to contain multiple hemes, one of which may be shared between both subunits (Dinauer et al., 1990)(Quinn et al., 1992). Recent studies suggest that the phagocyte cytochrome is actually a flavocytochrome b, containing the NADPH binding site and flavin (Rostrosen et al., 1992)(Segal et al., 1992) and is capable of producing O_2^- *in vitro* even in the absence of all other NADPH oxidase components (Koshkin et al., 1994). However, in the intact cell and in the cell-free reconstitution system, the cytosolic components of the oxidase are absolutely required for NADPH oxidase activity.

The absolute requirement for flavocytochrome b in the

NADPH oxidase is demonstrated by the inherited disorder known as chronic granulomatous disease (CGD) (Smith et al., 1991)(Dinauer et al., 1993). Phagocytes from patients with CGD are unable to produce O_2^-, and defects in the genes for gp91-*phox*, p22-*phox*, p47-*phox*, and p67-*phox* have been shown to be responsible for the various types of this syndrome (Smith et al., 1991)(Dinauer et al., 1993). The most common form of CGD results from a defect in the X-chromosome and results in the absence of gp91-*phox* (Dinauer et al., 1987)(Teahan et al., 1987) with a secondary loss of p22-*phox* (Segal, 1987)(Parkos et al., 1989). Studies of flavocytochrome b-deficient CGD neutrophils have also shown that neither p47-*phox* or p67-*phox* translocate to the plasma membrane when the cells are activated (Heyworth et al., 1991). Thus, the lack of translocation of cytosolic oxidase components in flavocytochrome-deficient CGD cells suggested that an intracellular domain of p22-*phox* and/or gp91-*phox* interacts with cytosolic NADPH oxidase components. This interaction is supported by studies of Kleinberg et al (Kleinberg et al., 1990)(Kleinberg et al., 1992) who found that synthetic gp91-*phox* carboxyl-terminal peptides inhibited superoxide generation by blocking assembly of the oxidase, and active peptides were found to disrupt the specific interaction between flavocytochrome b and p47-*phox* (Kleinberg et al., 1990)(Kleinberg et al., 1992)(Nakanishi et al., 1992). Thus, peptides with amino acid sequence

peptides representing the carboxyl-terminus of gp91-*phox* appear to compete with gp91-*phox* for binding to p47-*phox*. When the active peptides are bound to p47-*phox*, the peptides presumably closely mimic the active conformation of the carboxyl tail of intact gp91-*phox* when bound to p47-*phox*.

In the present studies, we use two-dimensional transferred NOESY NMR techniques to study the conformation of a synthetic bioactive gp91-*phox* carboxyl terminal peptide ^{552}SNSESGPRGVHFIFNKEN^{569} when bound to p47-*phox*. The conformation of the peptide induced by binding to recombinant p47-*phox* contains an extended open turn and the NMR experiments provided evidence for the interaction of certain key peptide residues with p47-*phox*.

Materials and Methods

Anti-p47-*phox* (B-3) (Volpp et al., 1989) was a kind gift of Drs. Robert A Clark and William Nauseef (University of Iowa).

Purified human neutrophils, isolated as previously described (Quinn et al., 1989), were treated with 3 mM diisopropylfluorophosphate (DFP) for 15 minutes at 4°C to inactivate serine proteases. The cells were washed; resuspended at 5 x10^8 cells/ml in 150 mM NaCl, 1 mM PMSF, 15 µg/ml leupeptin, 10 µg/ml chymostatin, 1 mM EGTA, and 10 mM Hepes (pH 7.0); and disrupted N_2 cavitation at 400 psi for 15 minutes at 4°C with slow stirring. Membrane and cytosolic fractions were prepared from the cavitate by

sequential centrifugation as described by Fujita et al. (Fujita et al., 1987). The resulting membranes and cytosol were aliquotted and stored at -70°C. Protein concentration was determined using the Pierce BCA assay (Pierce, Rockford, IL).

NADPH oxidase activity was measured in a cell-free system as described previously (Quinn et al., 1992). The reaction mixture containing 100 μM cytochrome c, 160 μM NADPH, 10 μM FAD, 1.5 x 10^6 cell equivalents of membrane, 10^7 cell equivalents of cytosol, and buffer (100 mM Kcl, 3 MM NaCl, 3.5 mM $MgCl_2$, 10 mM Pipes, pH 6.7) in a total volume of 1 ml was incubated for 3 minutes followed by the addition of 40 μM SDS. The reference cuvette also contained 45 μg/ml superoxide dismutase. The rate of O_2^- production was measured continuously by following superoxide inhibit superoxide reduction using a extinction coefficient of 18.5 mM^{-1} cm^{-1}.

To analyse peptide inhibition, the peptide was added to the reaction mixture before the 3 minute incubation. Peptide added after activation of the system with SDS failed to inhibit oxidase activity (not shown), consistent with the results of Kleinberg et al. (Kleinberg et al., 1990).

The gp91-*phox* carboxyl terminal peptide, ^{552}SNSESGPRGVHFIFNKEN^{569}, was synthesized on a Milligen Model 9050 peptide synthesizer using F-moc chemistry. Fmoc-Asn(Trt)-O PAC resin (Milligen/Biosearch) was used to

initiate the synthesis and the completed peptide was cleaved from the resin using 97% TFA/1% water/1.5% ethanedithiol/.5% thioanisole. The crude peptide was isolated by ether precipitation and purified by semi-preparative reverse phase HPLC in an acetonitrile/H_2O 2mM HCl gradient. The purity of the peptide was determined to be > 95% by electrospray mass spectrometry.

This peptide inhibited NADPH oxidase activity in a cell-free reconstitution assay system containing neutrophil membranes that was cytosol and activated with SDS (see above for assay conditions). This peptide also inhibited NADPH oxidase activity in electroporated whole neutrophil cells at <100 μM peptide (O. Dorseuiell, unpublished observation). Thus, our synthetic peptide preparation exhibits biological activity and has very similar characteristics to those described for this gp91-*phox* carboxyl-terminal peptide by Kleinberg et al (Kleinberg et al., 1990)(Kleinberg et al., 1992). The effective inhibitory concentrations of our peptide were higher than that previously reported (Kleinberg et al., 1992) and we attributed this to differences in assay conditions [e.g., different activators (SDS vs. arachidonic acid), different membrane preparations (whole membranes vs. detergent extracts of membranes), and assay volume (1 ml cuvette assay vs. microtiter plate assay)].

Recombinant p47-*phox* was produced in recombinant baculovirus-infected Spodoptera frugiperda (Sf9) cells

following the methods of Leto et al (Leto et al., 1991). Briefly, 300-400 ml suspension cultures of Sf9 cells were infected with recombinant baculovirus at a multiplicity of 5-10 plaque forming units/cell. The cells were seeded initially in TNM-FH + 10% fetal bovine serum at 2.5 X 10^5 cells/ml, and infections were performed when cell density reached 1-2.5 X 10^6 cells / ml. The cells were concentrated by centrifugation at 1000 x g 10 min. in a Beckman J6B centrifuge. Viral inoculum (7.8 X 10^7 plaque forming units/ml) was adsorbed to the Sf9 cells at 10^7 cells/ml for ~ 1 hour, and the cells with virus were transferred back into the flask containing the original, clarified medium with penicillin-streptomycin, 1:100 pluronic acid (F-68), and incubated at 27° C for 72-80 hours. The cells were pelleted at 2000 x g, washed with phosphate-buffered saline, repelleted, and processed immediately or frozen at -85° C for future processing as described below.

Cells were pelleted and resuspended in 50 mM KCl, 3 mM NaCl, 2 mM $MgCl_2$, 1 mM PMSF, 0.1 mM dithiothreitol, and 5 mM PIPES, pH 7.5, treated with 5 mM diisopropyl fluorophosphate for 15 minutes, disrupted by nitrogen cavitation (400 p.s.i. for 15 minutes), and centrifuged at 1500 x g for 10 minutes. The supernatant was recentrifuged at 1000,000xg for 1 hour and the supernatant was collected and diluted to 2.5 volumes with distilled H_2O. The diluted supernatants were loaded

onto a 5 ml Econo-Pac Q column that had been equilibrated with 5 mM KH_2PO_4, pH 7.0. The column was washed with 5mM KH_2PO_4 until the absorbance of the elutant stabilized at baseline, and bound proteins were eluted with an 80 ml gradient of 0-0.3 M NaCl in 5 mM KH_2PO_4, pH 7.0 (0.66 ml/min). Fractions (1.0 ms) were collected and analyzed for protein content using the Pierce BCA assay. Purity was assessed by SDS-PAGE and Western blotting with anti-p47-*phox* antibodies which confirmed the identity of the purified protein. The recombinant p47-*phox* was found to be active in supporting reconstituting NADPH oxidase activity in a cell-free reconstitution system (data not shown). Typically 8-10 mg of purified p47-*phox* was present in the peak fractions.

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out at 20°C using 7-18% gradient gels as described previously (Parkos et al., 1987). Electrophoretic mobility of the proteins in the samples was compared with the mobility of prestained standard proteins (BRL, Bethesda, MD). Proteins were visualized on the gels by staining with 0.1% Coomassie blue G in 50% methanol/10% acetic acid.

Western blotting was performed as described previously (Quinn et al., 1989). Transfers were blotted with 1μg/ml rabbit IgG for 3 hr at 20°C, followed by alkaline phosphatase conjugated goat anti-rabbit IgG secondary antibodies for 1 hr at 20°C, and developed using an alkaline phosphatase development kit (Kirkegaard & Perry

Laboratories, Gaitherburg, MD).

NMR spectra of the free gp91-*phox* carboxyl terminal peptide (^{552}SNSESGPRGVHFIFNKEN^{569}) was recorded using 2.5 mM samples of peptide dissolved in 450 µl of 50mM sodium phosphate buffer in H_2O and 50 µl of D_2O adjusted to a pH of 6.5.

Peak fractions of p47-*phox* were combined and concentrated to ~16 mg/ml using Amicon Centricon 30 microconcentrators. For studies of the bound peptide, 2.0 mgs of gp91-*phox* carboxyl terminal peptide in 0.25 ml of 50 mM sodium phosphate buffer pH=7.2 was incrementally added to a solution of 5.0 mgs p47-*phox* in 0.20 ml of 50 mM sodium phosphate buffer pH=6.5 while monitoring the pH and frequently adjusting it to 6.5. The final concentration of the peptide was 2.0 mM and that of the p47-*phox* 200 uM. A major experimental difficulty was met in the ability to maintain the solubility of the peptide/protein complex over times long enough for NMR studies. Many samples that initially were fully soluble slowly precipitated over time.

NMR spectra were measured on a Bruker AM 500 MHz-spectrometer with time proportional quadrature phase detection (Redfield et al., 1975). TOCSY spectra (Bax et al., 1985) were collected at 1°C with a MLEV-17 spin lock time of 100 ms. NOESY spectra (Jeener et al., 1979) were collected with mixing times of 75, 100, 125, 150, 200, and 300 ms that were randomly varied by 10% in order to reduce

zero-quantum coherence transfer (Macura et al., 1981).

Two-dimensional NMR spectra were processed on SGI Iris workstations using FELIX macros developed in our lab at MSU by Christophe Lambert. Apodization typically applied to 2D spectra were 60 degree shifted sine bell window functions in t_1 and t_2 with zero filling to 2048 complex points. Spectral widths of 5263 Hz were used for all spectra. Polynomial baseline corrections of the third order were used in t_2 for both the NOESY and TOCSY. The NOESY spectra used 256 points in t_1, the transmitter phase and preacquisition delay were finely adjusted to avoid any requirement for phase adjustment of the spectra during processing in t_2 which greatly improved baseline flatness (Marion et al., 1988). The D0 delay in t_1 was adjusted to be equal to IN/2=47.5 usec minus an allowance of $4\tau/\pi$ for the finite τ lengths of the hard pulses (Bax et al., 1991) and this allowed the phase in t_1 to be set to PH0=45 and PH1=-90 with little or no further phase correction. All spectra were acquired with the carrier on the water resonance which was suppressed with coherent decoupler irradiation in the case of the TOCSY and with jump and return read pulses in the NOESY (Plateau et al., 1982). A $1/\sin(\Delta\omega\ D2)$ correction was applied to the NOESY intensity to accurately compensate for the intensity distortion due to the nonuniform jump and return excitation. (Lambert et Dratz, unpublished work). The jump and return excitation correction retrieved

intensity loss which is especially important in the alpha proton region close to water. All spectra reported were collected at 1°C and referenced to TSP, 3(trimethylsilyl)propionic 2,2,3,3-d4 acid sodium salt.

^{1}H NMR $T_{1\rho}$ relaxation experiments were carried out using the pulse sequence: Predelay-90_ϕ-$(SL_\psi)_n$-acq, receiver phase ϕ, where the first pulse (90_ϕ) is from the transmitter RF and the spin lock pulse (SL_ψ) is from the decoupler RF. On our spectrometer there was a 90° hardware phase shift between the transmitter and decoupler pulses. The phase cycles of the (90_ϕ) and (SL_ψ) were (x -x -x x y -y -y y) and and (x -x x -x y -y y -y) respectively. The duration of the spin lock time was set from 1 to 200 ms by appropriate looping a 500 us SL_ψ. The $T1_\rho$ experiment was separately carried out on the 6Gly-αH and the 15Asn-NH side chain proton resonances, both of which exhibited clear changes in chemical shift between the free and bound states of the peptide. The RF carrier, ω was set on either the 6Gly-αH and 15Asn-NH resonance during spin lock and acquisition and was hopped to the H_2O resonance during the relaxation predelay where weak phase coherent irradiation was used for solvent suppression. $T_{1\rho}$ was measured as a function of spin lock power from 250 to 21,000 Hz which was directly calibrated by 90/180° pulse measurements at the higher spinlock powers (21,000 to 2400 Hz). Peak intensities were measured using the difference in the height of the peak and

of the peak and a cubic spline corrected baseline. T_1 and T_2 relaxation experiments on the bound and free gp91-*phox* carboxyl terminal peptide were carried out with inversion recovery and Carr-Purcell-Meiboom-Gill spin echo sequences, respectively using presaturation for solvent suppression.

NOESY cross peak intensities were divided by a resolved single proton diagonal area to correct for cross peak intensity loss due to T_1 relaxation and to provide fractional NOE values. The fractional NOE values were first converted to approximate proton distances from the bound NOESY at a mix time of 200 ms, using the isolated two spin approximation and the sidechain amino protons of Asn (1.8 Å) as a reference intensity.

Peak volumes were obtained by manually selecting boxes around the crosspeaks and using a Felix utility to sum the points in the box. A total of 126 NOE cross-peaks were identified in which there were 98 i,i and 28 i,i+1 constraints. The fractional NOESY intensities free of the peptide were subtracted from the bound TRNOESY intensities of the gp91-*phox* carboxyl terminal peptide at a mixing time of 200 msec. This correction is done in order to calculate the TRNOE signal of the bound gp91-*phox* carboxyl terminal peptide since the observed TRNOESY intensities are comprised of the sum of contributions from the free and bound species of gp91-*phox* carboxyl terminal peptide. The NOESY distances were converted to a distance restraint file format using the

FELIX database software package (Biosym, 1993). Stereospecific assignments were not made and pseudo atom corrections (Wuthrich et al., 1983) were applied to the appropriate methylene distance constraints. All NOESY distance constraints initially used force constants of 100 kcal/(mol/Å^2). Chirality restraints were applied on all chiral atoms. Restraint on the omega dihedrals was applied with a force constant of 300 kcal/(mol/rad^2) to maintain bonds in the trans peptide configuration.

Molecular dynamics, simulated annealing and a variety of analysis routines in the Discover v2.8, InsightII v2.2 and NMRchitect v2.0 packages (Biosym, 1993) were used. A family of approximate bound structures of the gp91-*phox* carboxyl terminal peptide was produced by a simulated annealing protocol which involved 60,000 steps of dynamics from 800° K with cooling to 300° K and four rounds of steepest descent and four rounds of conjugate gradient minimization. Mardigras (Borgias et al., 1990), which provides a full relaxation matrix analysis of NMR NOESY intensities, was used along with the best fit approximate structures from simulated annealing to produce improved distance constraints. Methyl, β, γ, δ protons that were nondistinguishable were treated as pseudoatoms by Mardigras where the resulting distance was measured from the geometric center of the degenerate protons. Programs used for conversion of Biosym file formats to Mardigras nomenclature

were written in our laboratory at MSU by David Poole.

Results

The ^{1}H NMR spectra of the gp91-*phox* carboxyl terminal peptide was assigned by the combined use of TOCSY and NOESY 2D NMR spectra using standard approaches (Wuthrich, 1986). Residues such as 7Pro, 8Arg, 10Val, 13Ile, and 16Lys were easily identified by their spin system patterns in a TOCSY. Most of the amide to alpha ^{1}H crosspeaks were resolved in the fingerprint region except for overlapping 5SerNH-αH and 14PheNH-αH resonances. The 2AsnNH-α crosspeak was not observed seen in the TOCSY, but was seen in the NOESY. Sequential connectivity was readily accomplished with the aid of the fingerprint region and the NH-NH connectivity of the NOESY. The sequence from residues 1 to 18 can be very cleary traced. Table 11 shows the ^{1}H resonance assignments of all 18 spin systems of the gp91-*phox* carboxyl terminal peptide at 1°C. The simplest kinetic scheme for a ligand in exchange with a receptor system is:

$$L + E \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} EL$$

where L stands for the ligand and E for the receptor. The cross relaxation observed σ_{is}^{obs} protons (i,s) on a ligand in fast exchange with a receptor system is represented by

$$\sigma_{is}^{obs} = p_f \sigma_{is}^{f} + p_b \sigma_{is}^{b} \qquad 45$$

Table 11. ^{1}H resonance assignments of the gp91-*phox* carboxyl terminal peptide sequence at 1°C, pH 6.5.

residue	HN	Hα	Hβ	others	
Ser 1					
Asn 2	8.67		4.83	2.85, 2.90	γNH_2 7.11, 7.82
Ser 3	8.67		4.44	3.86, 3.92	
Glu 4	8.61		4.37	1.99, 2.13	γCH_2 2.28
Ser 5	8.48		4.52	3.90	
Gly 6	8.37		4.12		
Pro 7			4.40	2.01, 2.28	γCH_2 1.90
					δCH_2 3.64
Arg 8	8.72		4.35	1.78, 1.87	γCH_2 1.61, 1.66
					δCH_2 3.14
					NH 7.12, 7.28
Gly 9	8.52		3.96		
Val 10	8.16		4.03	1.97	γCH_3 0.80, 0.84
His 11	8.36		4.53	2.94	2H 8.14
					4H 7.11
Phe 12	8.65		4.71	3.06, 3.16	2,6H 7.13
					3,5H 7.15
Ile 13	8.14		4.07	1.69	γCH_2 1.08, 1.38
					γCH_3 0.80
					δCH_3 0.80
Phe 14	8.48		4.53	3.05	2,6H 7.30
					3,5H 7.36
Asn 15	8.50		4.63	2.66, 2.76	γNH_2 6.99, 7.72
Lys 16	8.42		4.24	1.77, 1.84	γCH_2 1.44
					δCH_2 1.71
					εCH_2
Glu 17	8.63		4.30	1.92, 2.10	γCH_2 2.28
Asn 18	8.18		4.44	2.69, 2.77	NH_2 6.93, 7.67

where p_f and p_b are mole fractions of free and bound ligand, σ^f and σ^b are the cross relaxation rates of the free and bound ligand (Landy and Rao., 1989). A key to the transferred NOESY effect is that the σ^{obs} can be dominated by the bound ligand cross relaxation even when the free ligand is in large excess, $\rho_b\sigma^b >> \rho_f\sigma^f$. In order for equation 45 to hold the exchange rate must be short compared to the reciprocal of the fastest cross relaxation rate of the ligand in the bound state ($1/\sigma^b$) (Landy and Rao., 1989). This condition can be expressed in terms of the exchange constant:

$$k_{-1} >> \max|\sigma^b| \qquad 46$$

Failure of this fast exchange condition makes the cross relaxation rates from a TRNOESY very difficult to interpret.

Determination of the exchange rate in ligand-protein systems in the past have been done by measurement of transverse relaxation rates (T_2), using the Carr Purcell Meiboom Gill (CPMG) spin echo amplitude as a function of the length of the refocussing pulse train (Gerig et al., 1975)(Dubois et al., 1992). An important difficulty with this approach is T_2 relaxation is effected by spin-spin scalar coupling. Single protons can be isolated in the ligand by selective deuteration to obviate this difficulty, but we were not able to accomplish this. $T_{1\rho}$ relaxation measurements do not suffer from spin-spin scalar coupling modulation. Since $T_{1\rho}$ relaxation experiments are spin locked on a transverse axis, B_0 inhomogeneity effects are also nullified (Farrar et al., 1971). These advantages make $T_{1\rho}$ relaxation experiments potentially useful as a method to measure

exchange rates. $T_{1\rho}$ relaxation experiments as a function of the strength of the applied spin locking RF field, ω_{SL} (Stilbs et al., 1978) have been recently applied to measure exchange rates in enzyme ligand system (Davis, Pearlman and London, in press).

Davis, Pearlman and London showed that the relaxation rates of the free ($R^{f}_{1\rho}$) and bound ($R^{b}_{1\rho}$) system, can be described as:

$$R_{1\rho}^{(i)}(\beta_{(i)}) = \frac{1}{T_{1\rho}}^{(i)} = \frac{\cos^2(\beta_i)}{T_1^{(i)}} + \frac{\sin^2(\beta_i)}{T_2^{(i)}} \qquad 47$$

where (i) stands for the free or bound state and:

$$\cos(\beta_i) + \frac{\Delta\omega_i}{(\Delta\omega_i^2 + \omega_{SL}^2)^{\frac{1}{2}}} \qquad 48$$

$$\sin(\beta_i) = \frac{\omega_{SL}}{(\Delta\omega_i^2 + \omega_{SL}^2)^{\frac{1}{2}}} \qquad 49$$

where $\Delta\omega_i$ is the offset frequency between the carrier and resonance frequency of the free or bound species of interest. An effective magnetic field is experienced by the protons which are not fully along the transverse axis, and these protons are subjected to both T_1 and T_2 processes. To account for the effects of exchange between the two environments equations 47-49 are substituted into

$$R_2^{ex} = \frac{p_f}{T_2^f} + \frac{p_b}{T_2^b} + p_f p_b \Delta\omega^2 G(k_{ex}) \qquad 50$$

where $\Delta\omega$ is the difference in resonance frequency between the free and bound species. $G(k_{ex})$ represents the exchange matrix and algebra leads to equation 51.

$$R_{1\rho}^{ex} = \rho_f[\frac{\cos^2(\beta_f)}{T_1^f} + \frac{\sin^2(\beta_f)}{T_2^f}] + \rho_b[\cos^2\frac{(\beta_b)}{T_1^b} + \sin^2\frac{(\beta_b)}{T_2^b}] \qquad 51$$

$$+ p_f p_b \sin^2(\beta)\Delta\omega^2[\frac{k_{ex}}{(k_{ex}^2 + \omega_{SL}^2)}] \qquad 51$$

where β is the angle of the quantized spins between the effective field γB_{eff} and the static field γB_0. This equation expresses the relaxation rate in the rotating frame, $R_{1\rho}$ as a weighted sum of T_1 and T_2 in the bound and free states and exchange effects. Equation 51 predicts curves of $R_{1\rho}$ as a function of ω_{SL} in which the amplitude of the curve largely depends on p_b and $\Delta\omega$, while the shape of the curve depends largely on k_{ex}.

The points in figure 29 with error bars are experimental $T_{1\rho}$ data, while the solid line represents the best fit of equation 51 to the data. Several parameters can be adjusted in fitting equation 51, but some may be fixed independently. The $1/T_{1f}$ and $1/T_{2f}$ were determined experimentally by standard NMR experiments on the free peptides. T_1^b can be determined from T_1^{obs} by the equation

$$\frac{1}{T_1^{abs}} = \frac{(1-p_b)}{T_1^f} + \frac{p_b}{T_1^b} \qquad 52$$

where $k_{-1} >> (1/T_1^b)$. In the same manner, T_2^b can be determined from $1/T_2^{obs}$ by the equation

$$\frac{1}{T_2^{abs}} = \frac{p_f}{T_2^f} + \frac{p_b}{T_2^b} \qquad 53$$

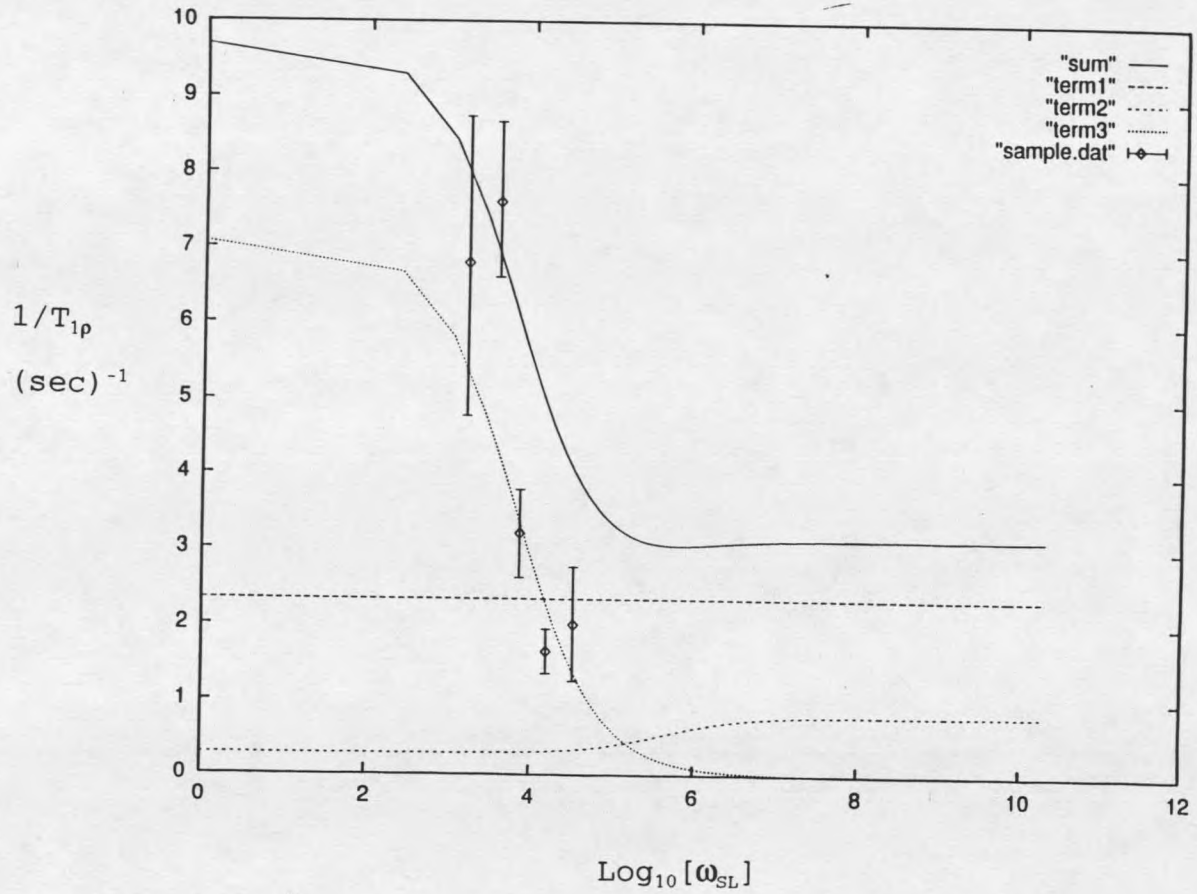
The first estimates of $\Delta\omega$ use to fit equation 40 were obtained by the shift differences observed for free and bound proton resonances. The fraction bound (ρ_b,) was varied between 0.005 and 0.05. The parameters used to fit figure 29 are shown in table 12.

Table 12. Parameters used to fit equation 40 for figure 29, where relaxation times are in (s^{-1}).

Proton	Temp (C)	T_1^b	T_2^b	T_1^f	T_2^f	$\Delta\omega$(Hz)	k_{-1}(s^{-1})
Asn NH (side)	1°	0.014	0.005	0.43	0.43	45	45

As can be seen in figure 29 the fitting of the experimental $T_{1\rho}$ data to equation 40 shows some discrepancy. Terms 1 and 2 in equation 40 caused an offset error for the fit at high spin lock powers where some experimental artifacts might be expected to occur because of heating and convection. In any case there is a clear effect of the third term (the exchange term) in equation 51, as the RF spin lock power is decreased. The break point in the third term indicates that exchange is taking place in the

Figure 29. $T_{1\rho}$ dependence on ω_{SL} for the 5Asn-NH sidechain proton of the gp91-*phox* carboxyl terminal peptide at $p_f:p_b$ of 10:1. The symbols used to denote the different terms in equation 51 can be seen in the top right of the figure. Experimental parameters for the fit of equation 51 are shown in table 12.



probable range of 30 to 50 (sec^{-1})(figure 29). The maximum cross relaxation that occurs in this system can be evaluated from the largest [NOE volume/diagonal volume]/(0.2 sec) and amounts to approximately 5 sec^{-1}. Therefore, the system investigated appears to be in fast exchange on the cross relaxation timescale.

The chemical shift and linewidth of the NMR spectra are dominated by the population of the free ligand $\rho_f >> \rho_b$. Conversely the TRNOE intensities are dominated by the bound ligand, since the cross relaxation between protons is very much faster when the ligand is bound to the slow moving protein than in the free rapidly tumbling ligand.

Figure 30 shows the upfield and downfield ^1H regions of free and bound gp91-*phox* carboxyl terminal peptide at a ligand to receptor concentration of about 10:1. As can be seen from the spectra, there is little chemical shift difference due to binding but there is evidence of some broadening of lines in the downfield amide and aromatic region in the bound state. The pH has been carefully adjusted to be equal in both samples. The sidechain amine of 2Asn, which in the free peptide is a single peak, becomes a doublet in the bound case and the ^1H resonance is shifted upfield ~7 Hz in the bound case. This shifting of the peptide ^1H resonance upon addition of the p47-*phox* protein and the observed line broadening is evidence of some type of chemical interaction between the peptide and protein in the system.

Many of the samples prepared formed precipitates soon after mixing p47-*phox* and the peptide. In cases where sample solubility could be maintained, the NOESY spectra were run rather quickly before precipitation could occur. Therefore a relatively

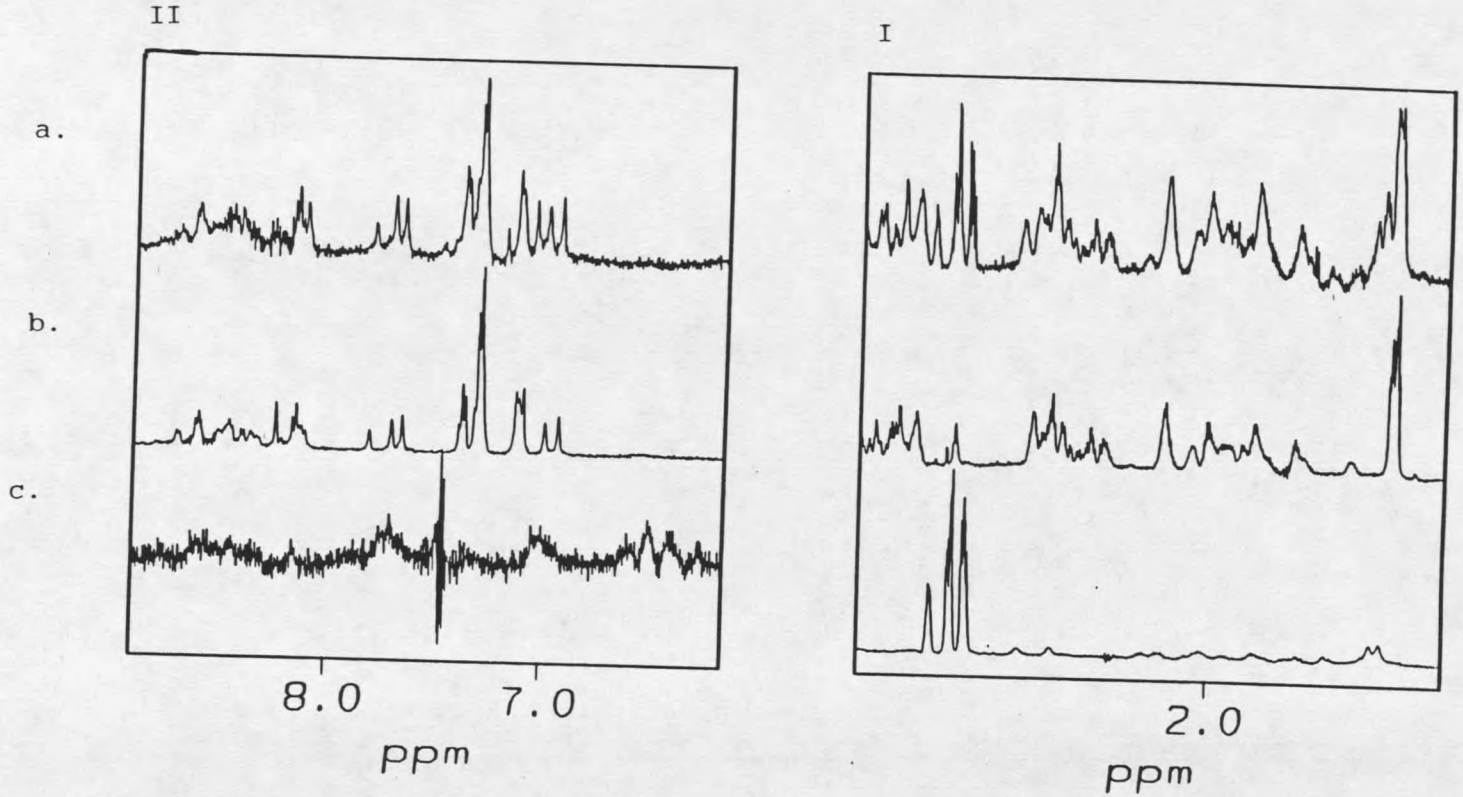
low number of scans and relatively few t_1 transients were used to keep the NMR measuring time below ca. eight hours at each mixing time. The NOESY spectra shown had modest signal/noise. The bioactive gp91-*phox* carboxyl terminal peptide in the presence of 0.1 mM p47-*phox* shows no new NOE crosspeaks but many of the NOESY crosspeaks either greatly increased or sometimes decreased in intensity. The weaker crosspeaks of the free gp91-*phox* carboxyl terminal peptide are not visible in the data shown but do appear at higher mixing times (data not shown). As can be seen from figure 30 many crosspeak intensities in the aliphatic sidechain region increased greatly in the presence of p47-*phox*, especially from residues 8Arg to 16Lys.

Volume integrals obtained from a 200 msec mixing time TR-NOESY were used to generate initial distance restraints as described in the materials and methods section. The distance restraints were subsequently refined with the full relaxation matrix refinement analysis program MARDIGRAS. MARDIGRAS generated distance constraints which are sensitive to the correlation time (τ_c). Assuming that the ligand bound to the receptor will have the same correlation time as the receptor, as a first approximation, we estimated the approximate correlation time of the p47-*phox* using the Stokes-Einstein equation

$$\tau_c = \frac{4\pi\rho r^3}{3kT} \qquad 54$$

where ρ is viscosity, r is the radius, k is the Boltzmann factor and T is the temperature. The correlation time of p47-*phox* if a sphere is 24 nsec. The effective τ_c of the gp91-*phox* carboxyl

Figure 30. 1D NMR upfield (I) and downfield (II) regions of a.) p_f:p_b=10:1 of p91-*phox* carboxyl. b.) free gp91-*phox* carboxyl terminal peptide. c.) p47-*phox*



145

terminal peptide in fast exchange is a weighted correlation. At a fraction bound of 0.1 the effective τ_c of the peptide would be ~2.4 nsec. Correlation times in this range were used in MARDIGRAS.

If the concentration of active p47-*phox* was less than 100% or the bound peptide experienced more local motion than the entire protein, the effective correlation time could be lower than 2.4 nsec. Three lower correlation times of 0.6, 0.8, and 1.0 nsec gave the best Q and R values (Withka et al., 1992) which are measures of agreement with the data in the MARDIGRAS refinements.

Initial distance constraints, derived from the isolated two spin approximation, were taken from the 2D NOESY of the p47-*phox* bound carboxyl terminal gp91-*phox* peptide that had been corrected for free peptide contributions. These NOESY distance constraints, along with ω dihedral constraints and chirality restraints on the chiral atoms of the peptide, were used in a simulated annealing protocol (SA) to generate 20 trial structures. From these 20 structures, 10 were discarded due to chirality violations. Of the remaining 10, 5 were picked which had the lowest relative energies and best distance agreements. These five structures were then used in an iterative MARDIGRAS refinement procedure. For each of the structures, MARDIGRAS was run with τ_c's of 0.6, 0.8, 1.0, 2.0, and 3.0 nsec. MARDIGRAS distances were then used in restrained molecular dynamics. Fifteen sets of structures from the molecular dynamics runs were analyzed for RMS distance agreement, ω dihedral violations and best relative energy. From each set of molecular dynamics

structures, one structure was chosen that had the best overall RMS distance agreement and lowest energy. Fifteen structures were then subjected to another round of MARDIGRAS refinement. Since the MARDIGRAS refinement gave the best Q and R data agreement at a τ_c of 0.6 nsec, the final round of MARDIGRAS was done at this correlation time. Again the resulting MARDIGRAS distances were used in constrained molecular dynamics and analyzed again for the best RMS distance agreements and lowest relative energy. The structure refinement was iterated once again with a third round of MARDIGRAS. The Q and R factors were similar in the second and third rounds of refinement and therefore it was concluded that further refinement would produce no further improvement.

The series of structures at τ_c of 0.6 nsec have the best Q and R factors from tables (13-15) as can be seen. Analysis of calculated experimental distance agreements also showed that the structures at τ_c of 0.6 and 0.8 nsec had fewer distance violations than the structures at a τ_c of 1.0 nsec or longer. There was a consistent distance violation over 0.5 Å at the 7Pro HG*-HD* restraint in all the structures. This can probably be explained by the appearance of ^{1}H buffer resonances occurring in the 3.6 ppm region where Pro ^{1}H resonances typically occur. The buffer peaks interfere with accurate measurement of the 7Pro:Hγ*-δ* cross peak intensity. Consistent ω dihedral violations over 4° were found for the 11His and 12Phe residues.

Table 13. Analysis of third round MARDIGRAS structure at a τ_c of 0.6 ns.

	Q	R	E(kcal)
A	0.053	0.111	417
B	0.099	0.202	489
C	0.083	0.180	411
D	0.071	0.154	410
E	0.058	0.123	421

Table 14. Analysis of third round MARDIGRAS structures at a τ_c of 0.8 ns.

	Q	R	E(kcal)
A	0.096	0.211	405
B	0.096	0.212	419
C	0.143	0.332	401
D	0.135	0.311	403

Table 15. Analysis of third round MARDIGRAS structures at a τ_c of 1.0 ns.

	Q	R	E(kcal)
A	-----	-----	----
B	0.144	0.336	468
C	0.205	0.516	485
D	0.162	0.388	517
E	0.156	0.369	371

Discussion

The 2D NMR studies of the free gp91-*phox* carboxyl terminal peptide showed no detectable secondary structure as judged from

the lack of characteristic NOE's for helical or turn structures (data not shown). Only $d_{\alpha N}$ (i,i+1) crosspeaks were observed which suggests that the predominant conformation is likely to be a largely extended structure (Dyson et al., 1991). The bound gp91-*phox* carboxyl terminal peptide showed no totally new crosspeaks compared to the free ligand but showed large increases in NOESY intensities, especially in the aliphatic sidechain region of the 2D NOESY spectra (figure 31&32). These results appear to indicate a generally extended conformation of gp91-*phox* carboxyl terminal peptide upon binding to gp47-*phox*.

The five third round MARDIGRAS-refined structure sets of the bound gp91-*phox* carboxyl terminal peptide were analyzed using utilities in NMRchitect. Figure 33 shows histograms for the five sets of structures. The top left shows the number of NOEs per residue while the other five histograms shows the NOE violation per residue above and below a 0.25 Å cutoff for the five sets of structures. For all five structures the total distance violations did not exceed 2 Å. As can be seen residues 7Pro, 8Arg, 12Phe, 13Ile, and 14Phe show the most NOE violations. Residue 7Pro NOE violation appears to be due to the buffer ^{1}H resonances from the buffer in gp47-*phox* which masks the proline Hγ and Hδ resonances. The largest NOE violation is the 8Arg-NH resonance to 9Gly-αH, which has a 0.8 Å upper violation which might indicate a different local motion in this region relative to the other positions of the peptide. NOE violations from 12Phe, 13Ile, and 14Phe are due to NH-αH, and NH-βH (i,i+1) NOEs.

Figure 34 shows the five best sets of clusters of structures of the bound gp91-*phox* carboxyl terminal peptide. All five

structures have clear differences in their conformations, particularly at their ends. However, all five sets of structures seem to have an open bend in the region of 5Ser to 9Gly. This sequence includes 2 glycines and a proline which are residues that typically favor Type I Beta or Type II Beta turns (Wilmot et al., 1988)(Richardson, 1981). But due to the lack of characteristic d_{NN} (i,i+2) and $d_{\alpha N}$ (i,i+2) crosspeaks, a tight turn does not appear to be present in the p47-*phox* bound structure. The most structurally defined sequence (8-14) of the gp91-*phox* carboxyl terminal peptide bound is thought to be the most important part of the peptide for interaction with p47-*phox*, since this is the minimum sequence capable of inhibiting NADPH oxidase activity (Kleinberg et al., 1992). Figure 35 shows the conformations of the five lowest energy frames superimposed on each other by their backbone atoms from residues 8Arg to 14Phe. The bound structure of the region between 8Arg and 14Phe shows a rather good definition of an open turn. As can be seen from the Ramachandran plots figure 36a-e there is general agreement in the phi/psi in all five structures between residues 8-14. The difference between the families of structures are located at a small number of phi/psi angles.

As can be seen in figures 31 and 32, many NOESY crosspeaks show very substantial increases in intensity in the bound state compared to the free peptide. The region of the data shown includes intramolecular interactions between protons on amino acid sidechains. The increases in NOESY crosspeak intensities of the peptide upon binding to p47-*phox* indicates a very substantial

Figure 31. 2D NMR aliphatic region of the bound gp91-*phox* carboxyl terminal peptide with p47-*phox*.

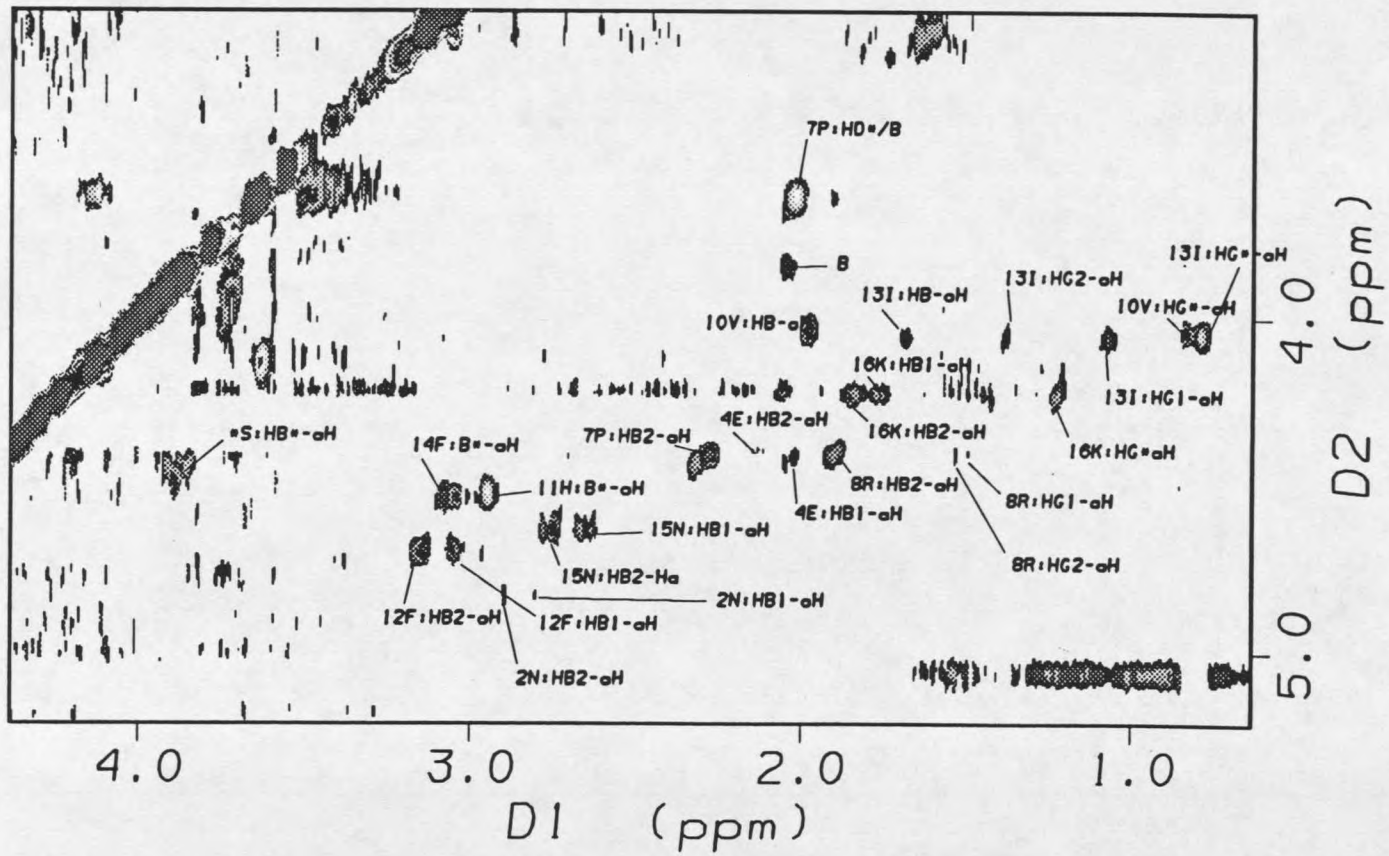


Figure 32. 2D NMR NOESY of the aliphatic region of the unbound gp91-*phox* carboxyl terminal peptide.

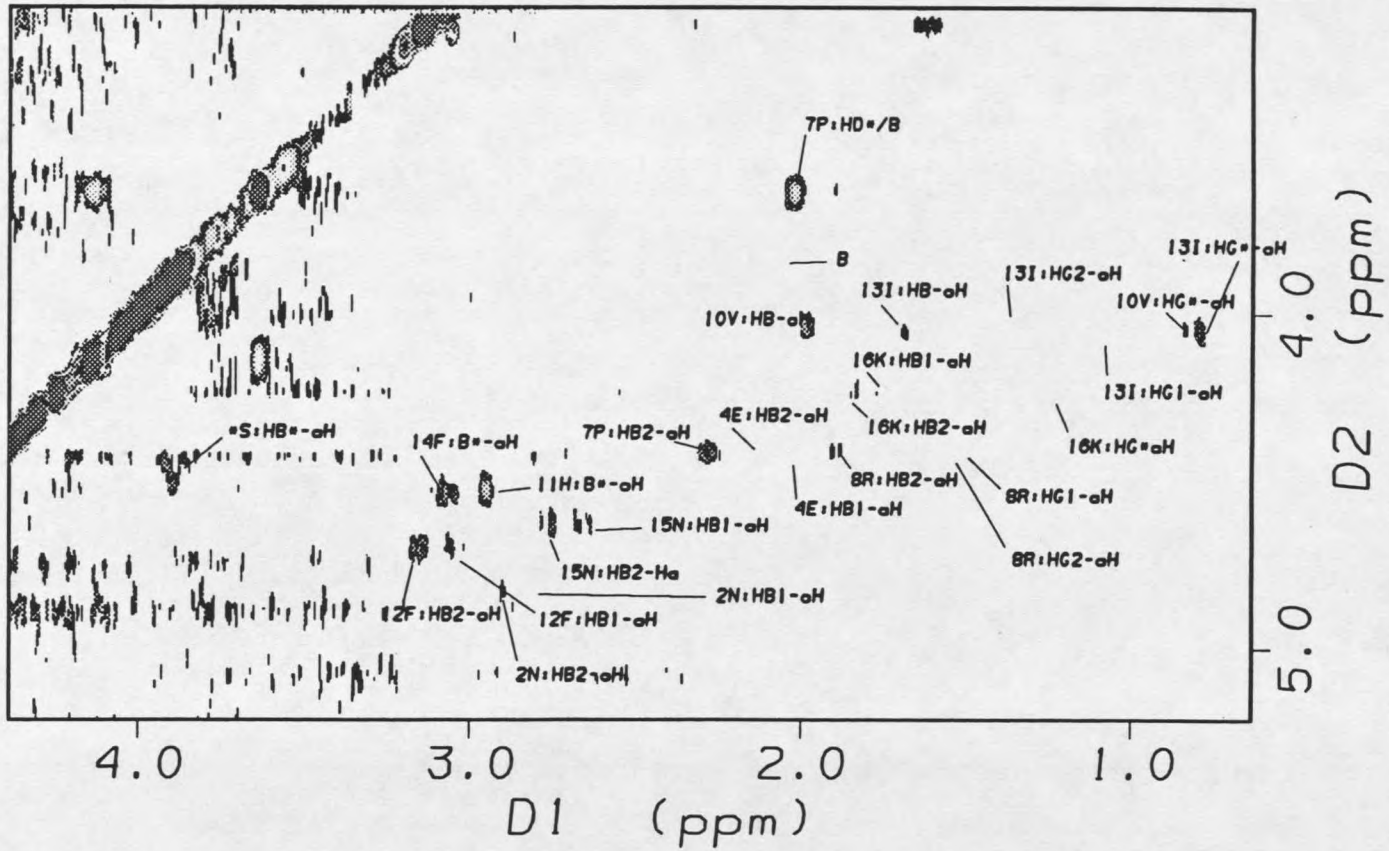


Figure 33. Histograms showing the number of NOE distance constraints per residue, and NOE violation per residue.

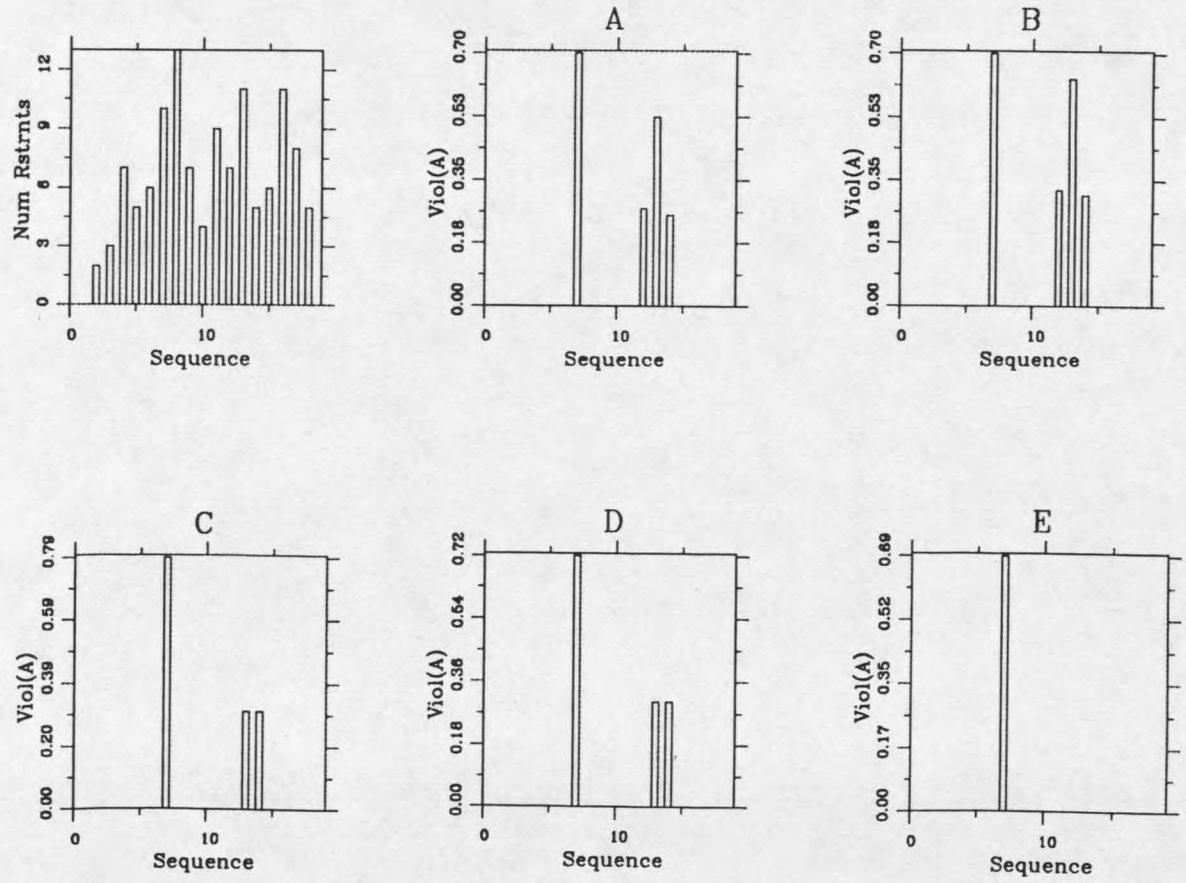


Figure 34. Five best structures from the clusters of the sets of the bound gp91-*phox* carboxyl terminal peptide.

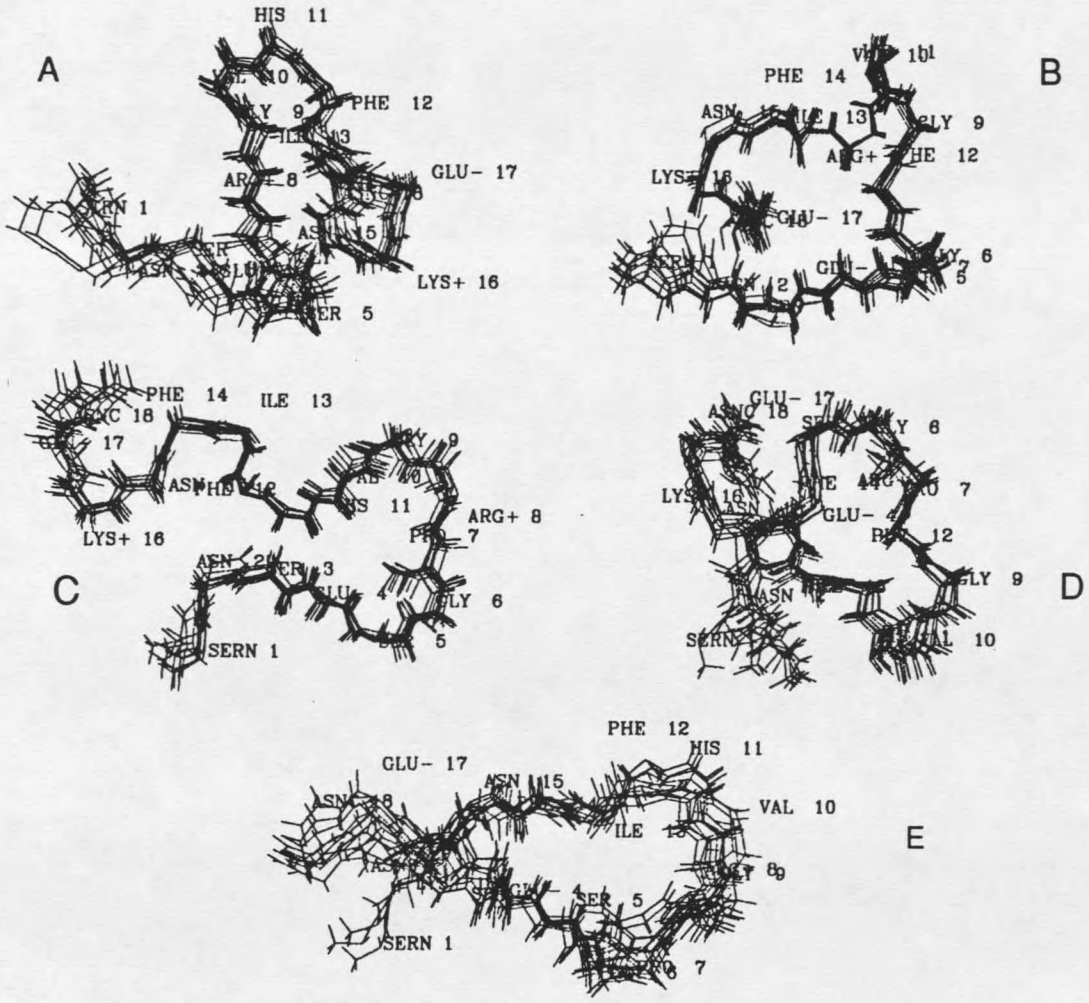
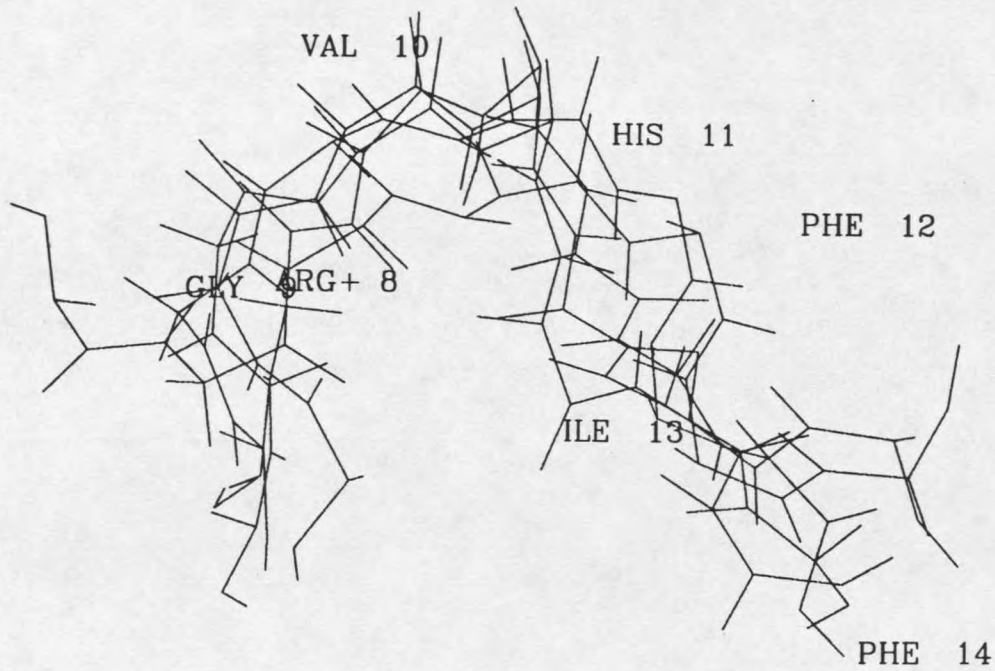


Figure 35. Structures superimposed from each of the fives sets from figure 34, showing the backbone atoms from resides 8Arg to 14Phe.



155

Figure 36. a. Ramachandran plots for residues 8Arg-14Phe in set a.

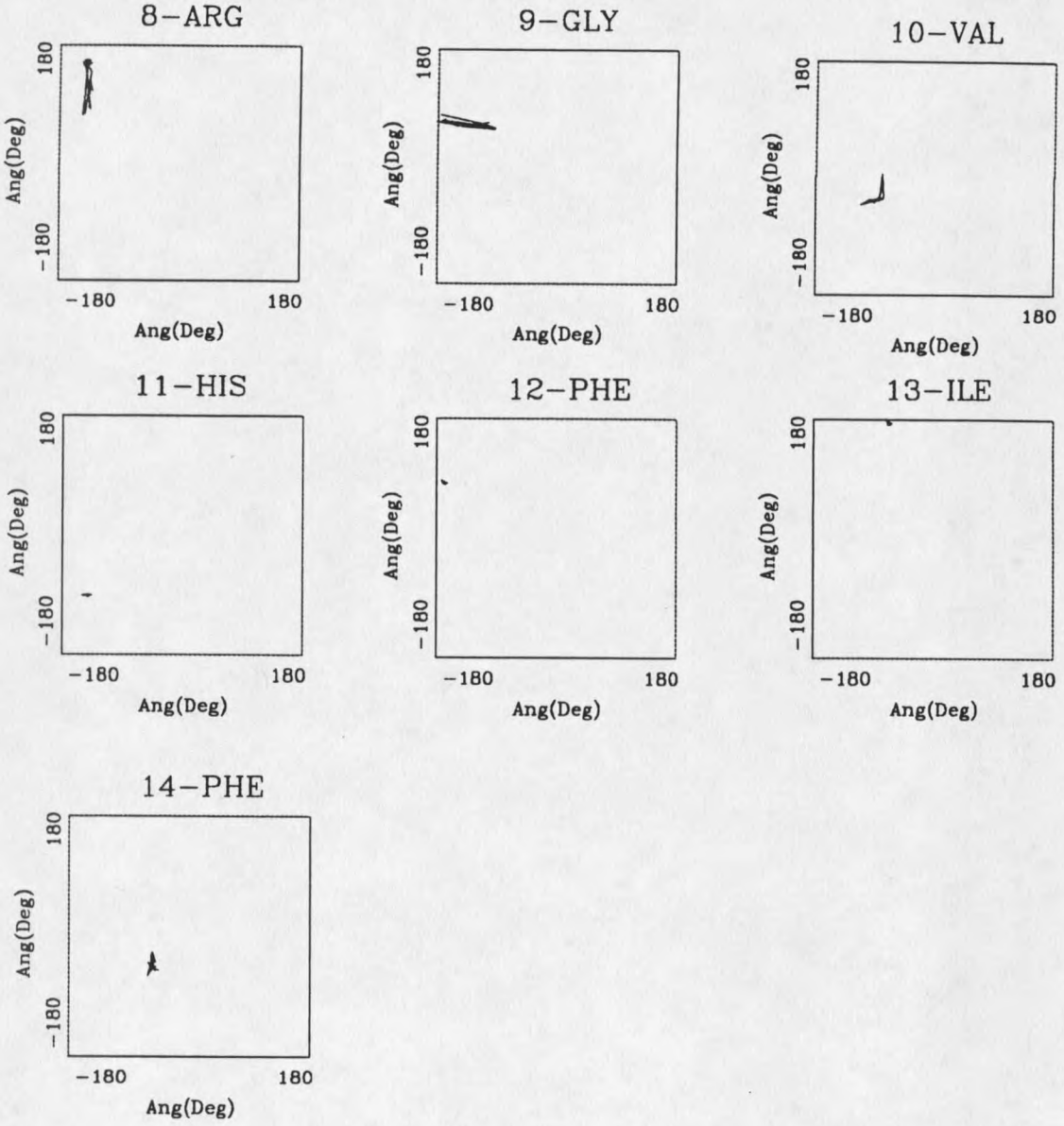


Figure 36 (continued). b. Ramachandron plots for residues 8Arg to 14Phe for set b.

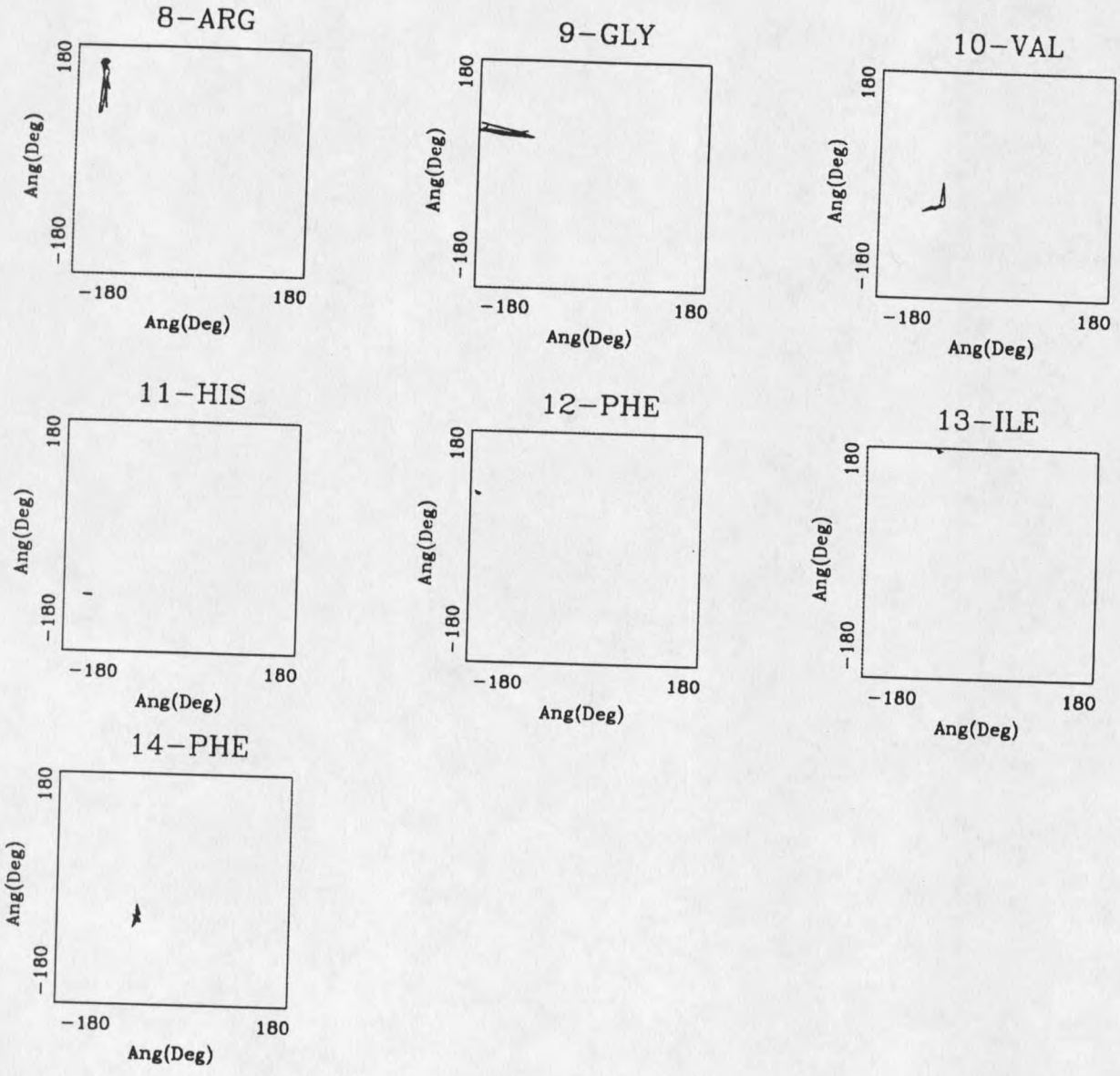


Figure 36 (continued). c. Ramachandran plots for residues 8Arg to 14Phe for set c.

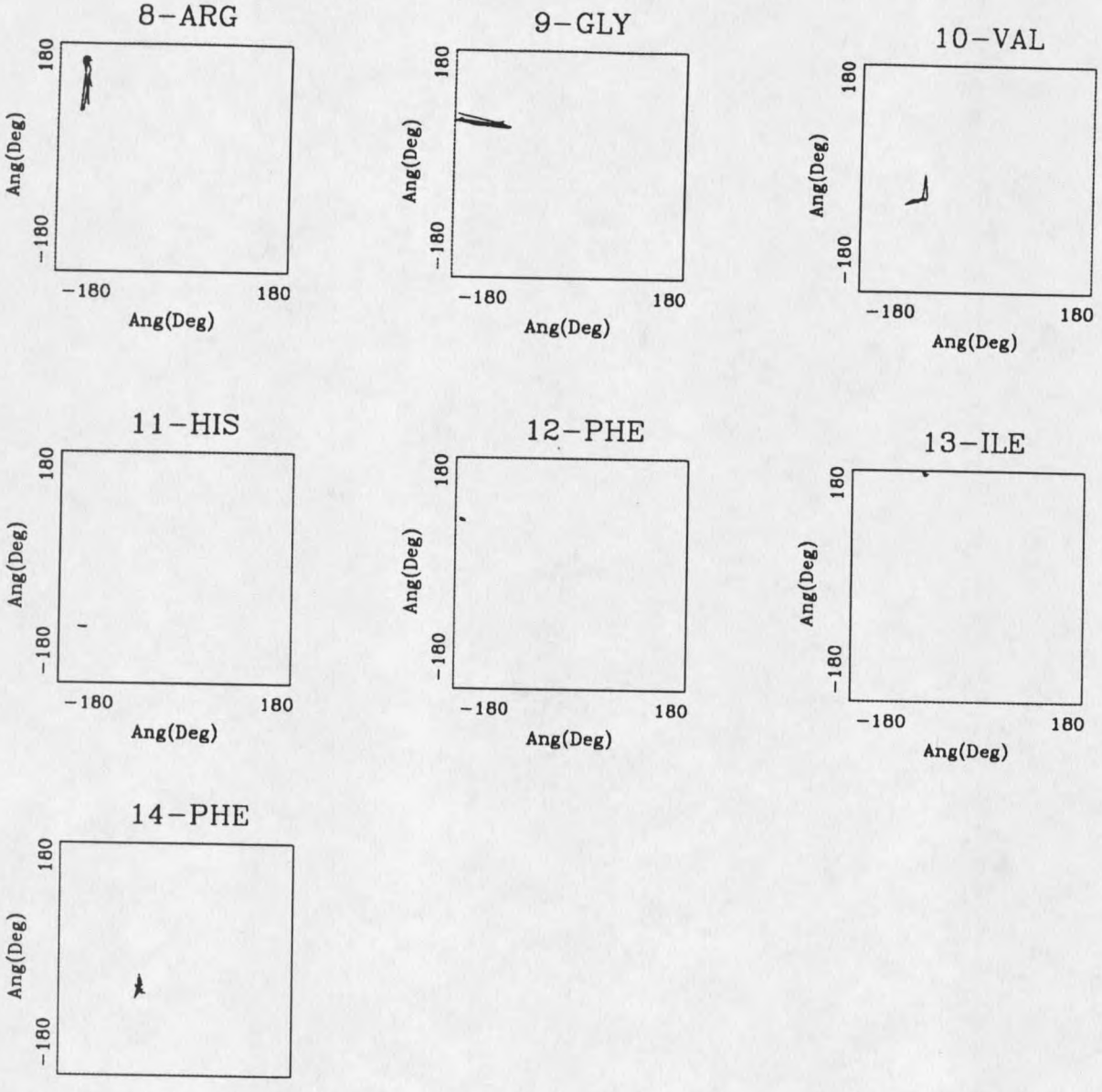


Figure 36 (continued). d. Ramachandran plots of residues 8Arg to 14Phe for set d.

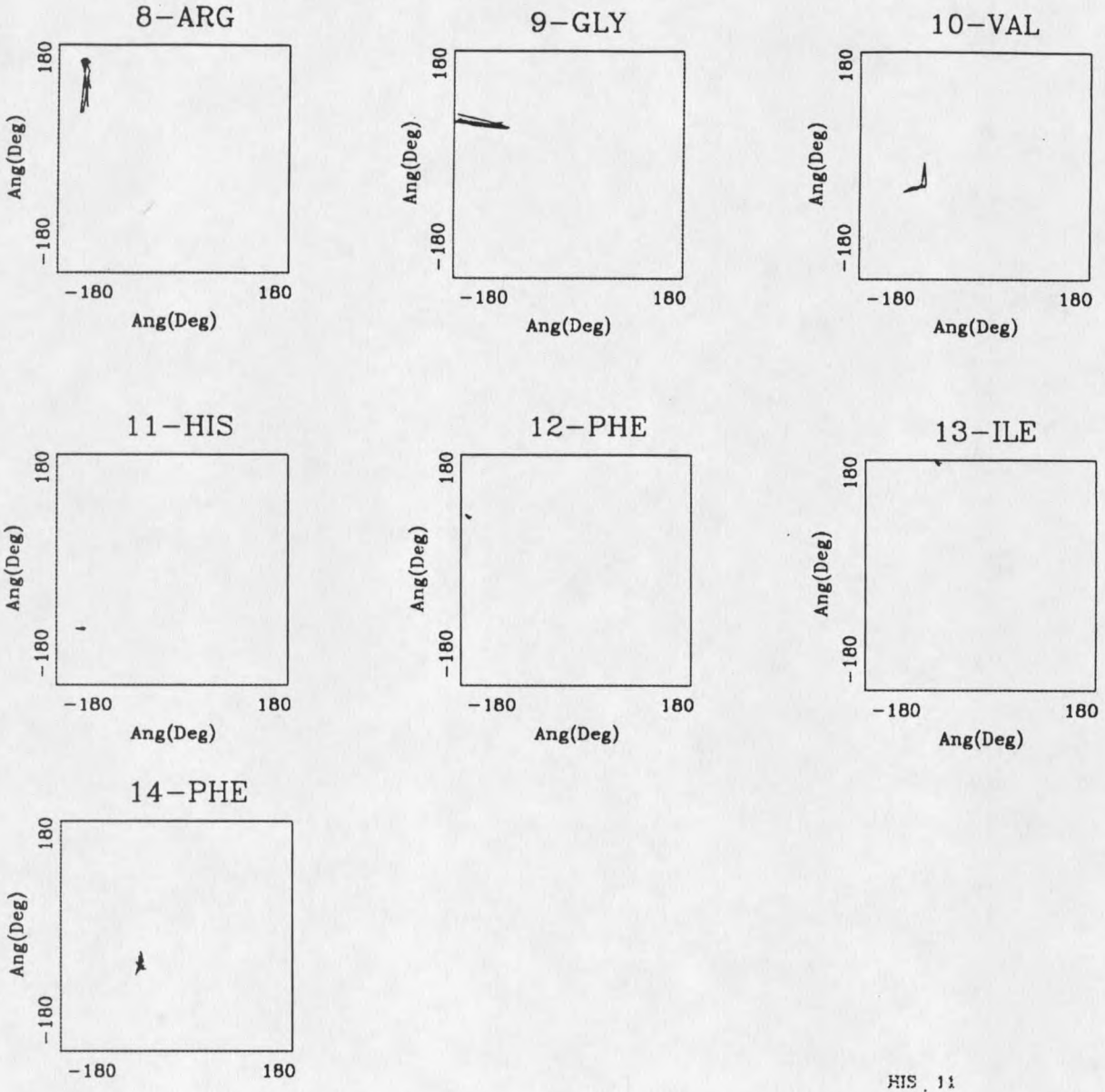
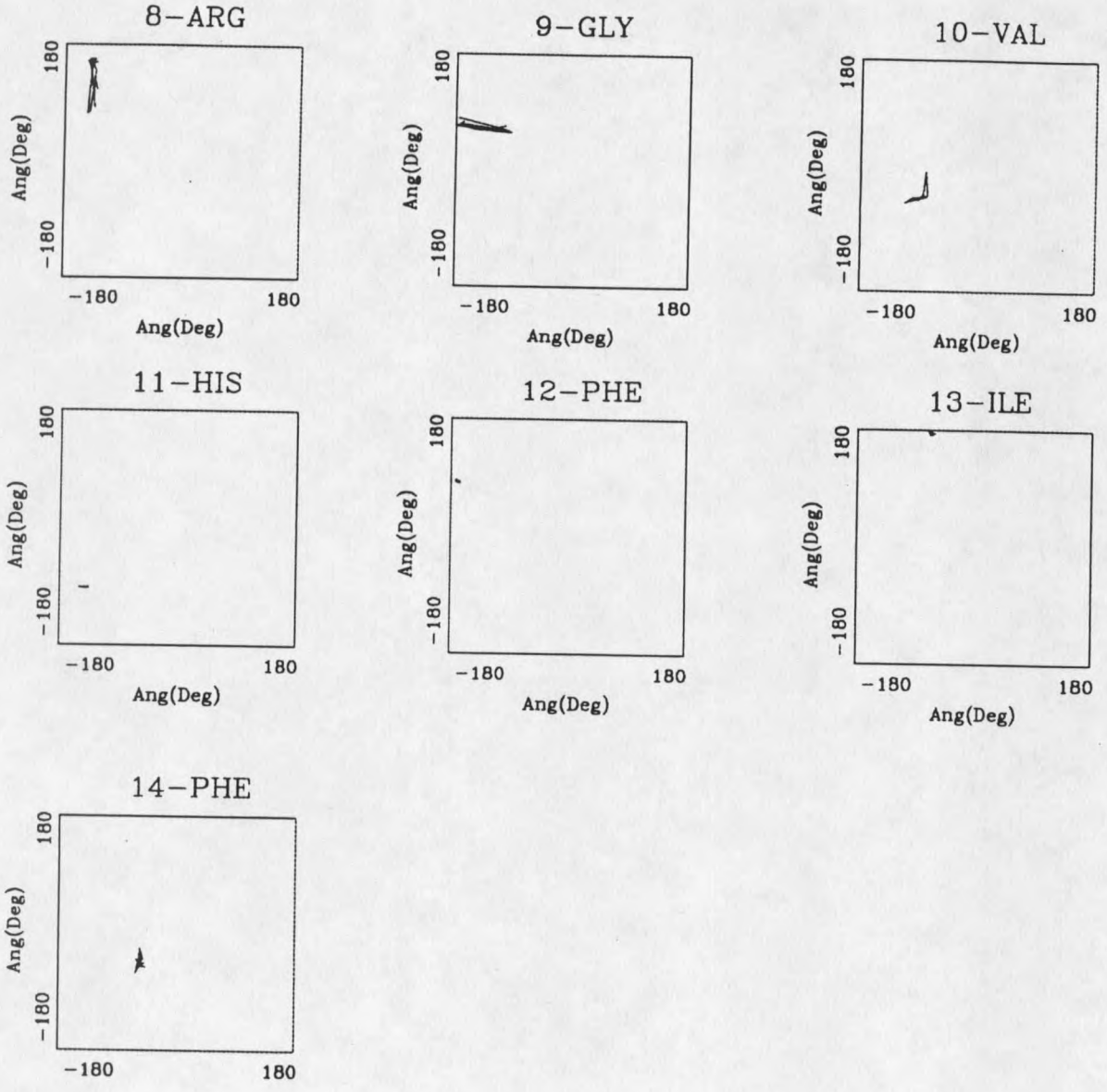


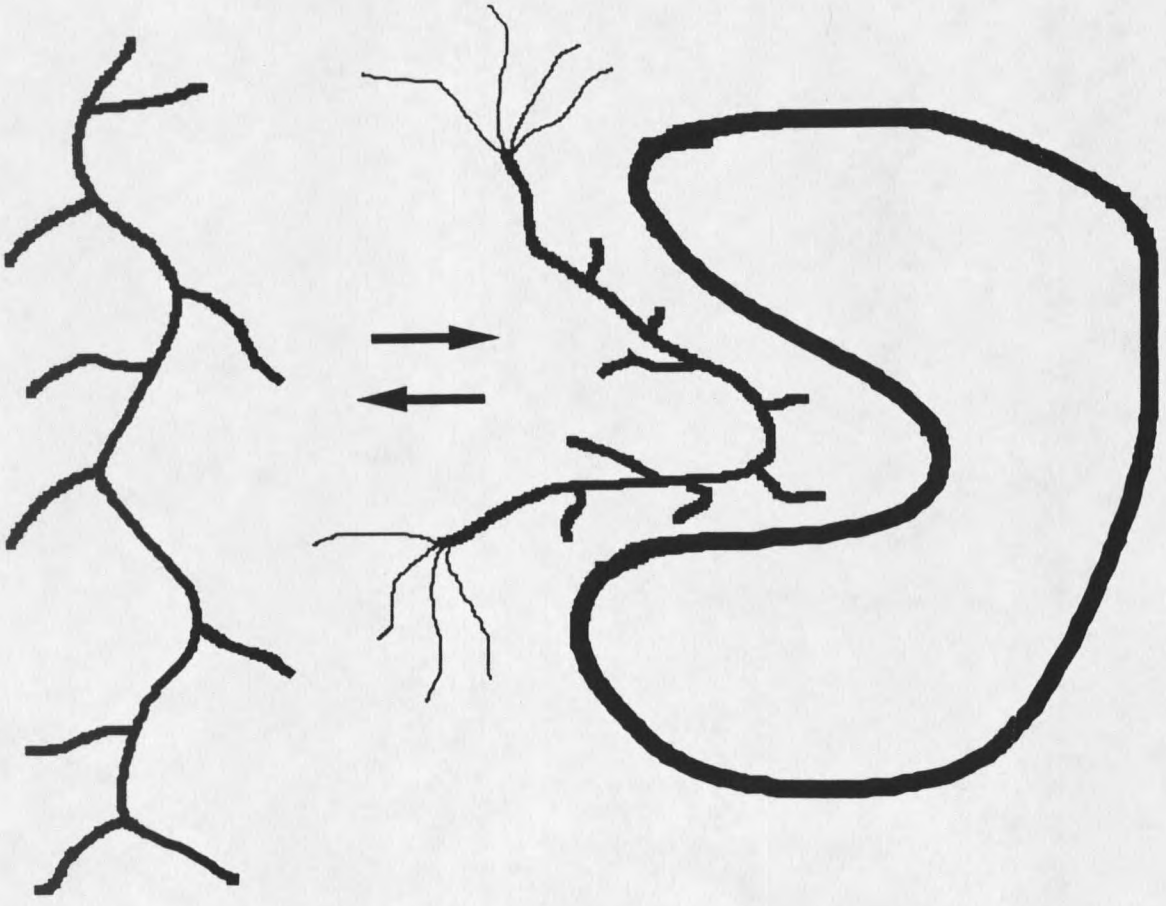
Figure 36 (continued). e. Ramachandran plots of residues 8Arg to 14Phe for set e.



immobilization of certain peptide sidechains by interaction with the p47-*phox* protein. The conformation of the ends of the peptide are not well defined by the TRNOESY experiments reported here. The model in figure 37 summarizes the current view of the conformation of the active peptide bound to p47-*phox*. The main chain from residues 8-14 appears to be in an open bend with close association of most of the peptide sidechains with the protein except 11 His.

There are several potential pitfalls in interpreting the TRNOESY data for bound ligand structures, primarily because of the possibility of nonspecific binding and spin diffusion influences of the protein ligand site. Spin diffusion (Kalk et al., 1976) is a problem in accurate distance determination in NOESY and specific pathway may be a problem in TRNOESY experiments (Zheng et al., 1993)(Nirmala et al., 1992). Intramolecular spin diffusion can be largely accounted for by complete relaxation matrix analysis as we have done using MARDIGRAS. Intermolecular spin diffusion between peptide and protein must also be considered. It has been shown in the fast exchange region, that the effective mixing time in a TRNOESY is equivalent to the experimental mixing time multiplied by the fraction of bound ligand (Landy et al., 1989)(Lippens et al., 1990). In the present 2D NMR NOESY experiment the ratio of free ligand to bound is somewhere between 10:1 and 20:1. With a mix time of 200 ms, spin diffusion is only effective for ~10-20 msec (the fraction bound x mixing time). In the case of the bound gp91-*phox* carboxyl terminal peptide relatively little influence of spin diffusion should occur in the time span in which it is

Figure 37. Model (cartoon representation) of the gp91-*phox* carboxyl terminal peptide binding to the p47-*phox* protein.



bound to gp47-*phox*. In the future more sophisticated methods of data analysis under development (Dratz personal communication) are expected to be able to more directly evaluate intermolecular effects.

Another factor that must be addressed in the interpretation of TRNOESY data is the possible effect of added protein on solution viscosity. Viscosity (Amodeo et al., 1991) can play a role in the change in the intensity of the NOESY crosspeaks of small peptides by affecting the correlation time which in turn has a direct effect on the rate of %NOESY buildup. Increased viscosity should also broaden resonances. Figure 30 shows the 1D NMR of the free and bound gp91-*phox* carboxyl terminal peptide and there is little broadening in the resonances in the 1D NMR which indicates that viscosity does not appear to be a large factor in this case. In the amide region between 8Arg and 6Gly there is a broad lump in the bound spectrum which likely is due to poorly resolved peaks from the more flexible portions of the p47-*phox* protein.

As mentioned above, one of the key factors in the successful use of TRNOESY in ligand protein systems is knowing that the chemical exchange rate is sufficiently fast. In our case the k_{off} seems to be between 30 and 50 sec^{-1} by the fitting of $T_{1\rho}$ to equation 40 and observing the inflection in term 3. This exchange rate appears to be in the intermediate region of kinetic exchange relative to the maximum difference in chemical shift between the bound and free states. However, with respect to processes such as T_1, and cross relaxation, the k_{off} is roughly a factor of 100 times faster than T_1, and 20 times greater than the

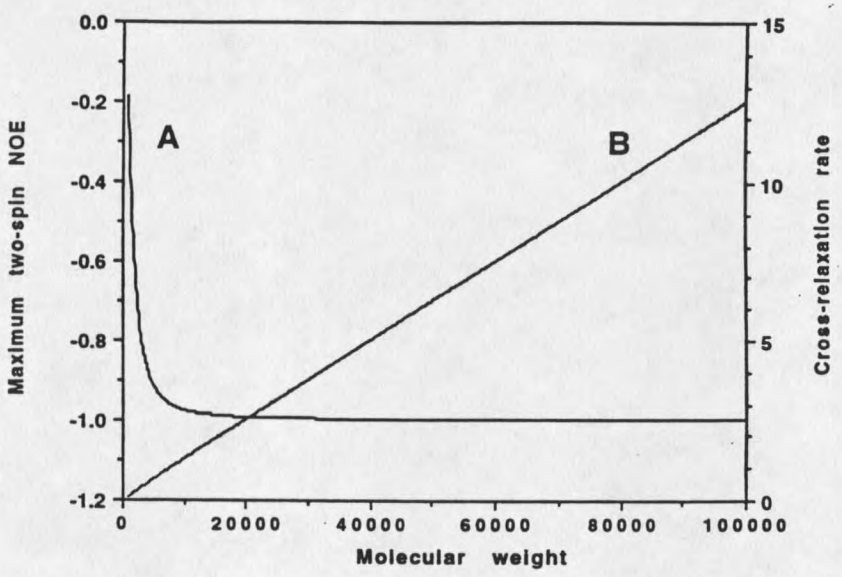
maximum cross relaxation rate of a ligand bound to a protein the size of p47-*phox* (figure 38) (Campbell et al., 1993). The fastest measured cross relaxation rate of the bound peptide in our experiments was 1.85 sec^{-1} which is 30 times slower than the rate of k_{ex}. So relative to the protein's cross relaxation and T_1 relaxation the exchange rate of the gp91-*phox* carboxyl terminal peptide is quite fast, therefore the exchange rate of the system is sufficiently fast to be interpreted in the fast exchange limit for TRNOESY.

The maximum observed cross relaxation rate for the bound peptide was found to be 1.8 sec^{-1} for a proton pair with a known short distance: the Ile γ_1-γ_2. If the peptide possessed no local motion when bound to the p47-*phox* protein the cross relaxation rate predicted for a 47,000 MW sphere and r_{ij}=1.8 Å would be ~3.5 sec^{-1}. The most likely explanation for this discrepancy may be the peptide is not rigidly bound to the protein and has local motion and an order parameter (Lipari and Szabo, 1982) of about 0.5. Alternatively it is possible that p47-*phox* is elongated and has a faster τ_c around some axes than a sphere which could give a smaller cross relaxation rate to a rigidly bound peptide.

Our NMR evidence indicates that the gp91-*phox* carboxyl terminal peptide binds to gp47-*phox* in an open bend with a generally extended conformation (figure 35&37). There are other cases in which a peptide ligand binds to a protein in an extended conformation for example the hexapeptide substrate to porcinic pancreatic elastase (Meyer et al., 1988); a dodecapeptide sequence of nicotinic acetylcholine receptor (nAchR) to a α-Bungarotoxin (Basus et al., 1993); and platelet receptor peptide to Bovine α-

Thrombin (Ni et al., 1992). There are several peptide/protein systems which show a structural change from an approximately

Figure 38. A.) maximum two-spin NOE as a function of molecular weight. B.) Cross relaxation rate between two protons at a distance of 2.5 Å as a function of molecular weight. Larmor frequency is 500 MHz (Campbell and Sykes, 1993).



random extended peptide in solution to a bound peptide with secondary structure; Troponin I peptide binds to skeletal Troponin C protein in an α-helix form (Campbell et al., 1991); the C-terminal peptide of hiruden bound to thrombin forms an α-helix (Ni et al., 1990); a calmodulin target peptide bound to calmodulin forms an α-helix (Ikura et al., 1992); Mastoparin peptide bound to a Giα protein in an α-helix conformation (Sukumar et al., 1992). Other protein bound peptides exhibit more complex shapes including beta bends (Dai, Ostheimer, Starky and Dratz, in progress).

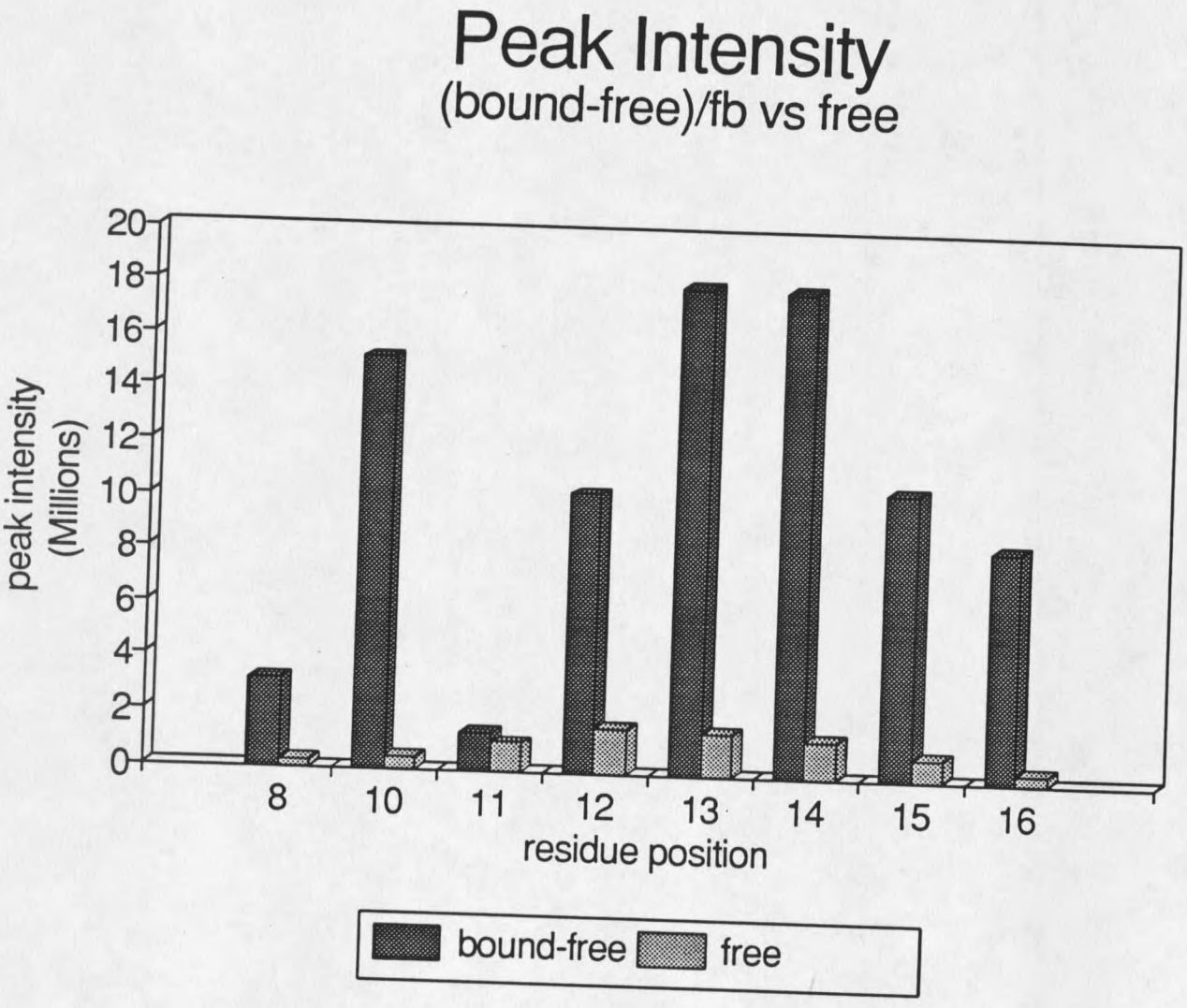
We have presented sets of the best fit structures of the

gp91-phox carboxy terminal peptide bound to p47-phox. When all of the best-fitting structures of the p47-bound gp91-*phox* carboxy terminal peptide are compared (figure 34) it seems clear that the structures of the ends of the p47-bound peptide differ and are not well defined. However, the bound structues of the center region of the peptide between residues 8 and 14 does appear to be fairly well defined (figure 35).

Figure 12 shows that large increases in the TRNOESY intensities of certain amino sidechains, relative to the free peptide NOESYs indicate that these side chains are relatively immobilized by binding to the p47-*phox*. Interestingly, the side chains between residues 8 and 16 of the peptide except for 11His residue sidechain appear to be interacting most strongly with p47-phox. There is no evidence on this point for residue 9 since the glycine side chain protons were not resolved. The relationship between the peptide and p47-*phox* is shown in the model in figure 37.

Kleinberg et al, have shown that the residues of the peptide that are most important for biological activity (inhibition of superoxide production) are between residues 8 and 14 and that extensions on either end of this region did not increase activity. Kleinberg et al carried out alanine scanning between residues 8 and 14. Residues 8Arg, 10Val, 12Phe, 13Ile and 14Phe all appear to be very important for activity, whereas 9Gly had modest effect and 11His could be replaced with Ala with little reduction in activity. Therefore residues that are most important for activity also show the strongest immobilization when interacting with p47-phox (figure 39).

Figure 39. Peak intensity plot of (bound-free)/fraction bound vs free from residues 8 to 16.



There are a number of improvements that could be made in these experiments to obtain more information on the strucure of the active interface between gp91-*phox* and p47-*phox*. The major limitation of the current experiments was the low solubility of the p47-*phox* protein-peptide complex under the pH and temperature conditions suitable for the NMR experiments. It may be possible to increase the solubility of the p47-*phox* protein-peptide complex with low concentrations of detergents which do not create backgound problems in the NMR. This direction is being successfully taken in another system under study in the laboratory where there are similar limitations in solubility of the target protein (Dai, Starkey and Dratz, personal communication). It would be useful to examine TRNOESY using control peptides which have very similar amino acid sequences but which have low or no activity to obtain controls for nonspecific interactions.

One control peptide that should be used is the reverse peptide. Other useful control peptides with closely related sequences but with very low activity such as Arg8Ala, Phe12Ala or Phe14Ala may be useful. Other blocking peptides that have recently been found by phage display can also be tested for specifically blocking the effects of the C-terminal gp91-*phox* carboxy terminal peptide. High affinity specific blockers would be useful if they exchange too slowly to show Tr-NOESY but may be able to block specific interactions with bioactive peptides that are in rapid exchange such as the C-terminal gp91-*phox* carboxy terminal peptide.

The approach outlined in the above chapter appears to have

great promise for determining the 3D atomic structures in regions of protein-protein interactions. Protein-protein interactions are at the heart of many biological signalling and adhesion mechanisms yet are usually not approachable by x-ray diffraction or full NMR structure determination.

REFERENCES

(1993) *Insight II Reference Guide*. Biosym Technologies Inc., San Diego

(1991) *Discover User Guide*. Biosym Technologies Inc., San Diego.

(1993a) *NMRchitect*. Biosym Technologies, San Diego.

(1993b) *Felix*. Biosym Technologies, San Diego.

Abo, A., Pick, E., Hall, A., Totty, N., Teahan, C.G. & Segal, A.W. (1991) *Nature 353*, 668-670.

Aime, S., Gobetto, R., Nano, R. & Santucci, E. (1987) *Inorg.Chim. Acta 129*, L23-L25.

Alcock, N.W., Tracy, V.M., and Waddington, J.C. (1976) *J.C.S. Dalton*, 2243-2246.

Amodeo, P., Motta, A., Picone, D., Saviano, G., Rancredi, T. & Temussi, P.A. (1991) *J.Magn.Reson. 95*, 201-207.

Anglister, J., Levy, R. & Scherf, T. (1989) *Biochem. 28*, 3360-3365.

Ariki, M. & Lanyi, J.K. (1986) *J.Biol.Chem. 261(18)*, 8167-8174.

Babior, B.M. (1987) *TIBS 12*, 241-243.

Badwey, J.A. & Karnovsky, M.L. (1980) *Annu.Rev.Biochem. 49*, 69-726.

Baggiolini, M., Boulay, F., Badwey, J.A. & Curnutte, J.T. (1993) *FASEB 7*, 1004-1010.

Balaram, P., Bothnerby, E. & Breslow, E. (1972a) *J.A.C.S 94*, 4017-4020.

Balaram, P., Bothnerby, A.A. & Dadok, J. (1972b) *J.A.C.S 94*, 4015-4017.

Baleja, J.D., Moult, J. & Sykes, B.D. (1990) *J.Magn.Reson. 87*, 375-384.

Barnett, B.L. & Uchtman, V.A. (1979) *Inorg.Chem. 18(10)*, 2674-2678.

Barsukov, I.L., Nolde, D.E., Lomize, A.L. & Arseniev, A.S. (1992) *Eur.J.Biochem. 206*, 665-672.

Basus, V.J., Song, G. & Hawrot, E. (1993) *Biochem. 32*, 12290-12298.

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

Bax, A., Ikura, M., Kay, L.E. & Zhu, G. (1991) *J.Magn.Reson. 91*, 174-178.

Beutler, E., Schneider, K.E. & Frincke, L.J. (1986a) *Informatics in Pathology 1*, 83-93.

Billeter, M., White, A.D., Braun, W., Huber, R. & Wuthrich, K. (1989) *J.Mol.Biol 206*, 677-689.

Bjorkman, P.J. (1987) *Nature 329*, 506-512.

Blair, D.F. & Berg, H.C. (1991a) *J.Mol.Biol. 221*, 1433-1442.

Boelens, R., Koning, T.M.G., VanDerMarel, G.A., VanBoom, J.H. & Kaptein, R. (1989) *J.Magn.Reson. 82*, 290-308.

Borgias, B.A. & James, T.L. (1988) *J.Magn.Reson. 79*, 493-512.

Borgias, B.A. & James, T.L. (1990) *J.Magn.Reson. 87*, 475-487.

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

Brown, J.H., Jardetzky, T.S., Gorga, J.C., Stern, L.J., Urban, R.G., Strominger, J.L. & Wiley, D.C. (1993) *Nature 364*, 33-39.

Butt, H.J., Wang, D.N., Hansma, P.K. & Kuhlbrandt, W. (1991a) *Ultramicroscopy. 36*, 307-318.

Campbell, A.P. & Sykes, B.D. (1991) *J. Magn. Reson. 93*. 77-93.

Campbell, A.P. & Sykes, B.D. (1991) *J.Mol.Biol. 222*, 405-421.

Campbell, A.P. & Sykes, B.D. (1993) *Annu.Rev.Biomol. Struct. 22*, 99-122.

Cao, Y., Varo, G., Chang, M., Ni, B.F., Needleman, R. & Lanyi, J.K. (1991) *Biochem. 30*, 10972-10979.

Chacko, V.P., Ganapathy, S. & Bryant, R.G. (1983) *J.A.C.S 105(16)*, 3491-3492.

Chang, C.H., Chen, J.G., Govindjee, R. & Ebrey, T. (1985) *Proc. Natl.Acad.Sci. 82*, 396-400.

Chang, C.H., Jonas, R., Melchiore, S., Govindjee, R. & Ebrey, T.G. (1986) *Biophys.J. 49*, 731-739.

Cheetham, A.K., Dobson, C.M., Grey, C.P. & Jakeman, R.J.B. (1987) *Nature 328*, 706-707.

Cheetham, J.C., Redfield, C., Griest, R.E., Raleigh, D.P.,

Dobson, C.M. & Rees, A.R. (1989) *Methods of Enzymology 203*, 202-228.

Chiu, W. (1993) *Ann.Rev.Biomol.Struct. 22*, 233-255.

Chou, P.Y. & Fasman, G.D. (1977) *J.Mol.Biol. 115*, 135-175.

Clark, R.A., Volpp, B.D., Leidal, K.G. & Nauseef, W.M. (1990a) *J. Clin.Invest. 85*, 714-721.

Clark, R.A. (1990b) *J.Infect.Dis. 161*, 1140-1147.

Clore, G.M. & Gronenborn, A.M. (1982) *J.Magn.Reson. 48*, 402-417.

Clore, G.M. & Gronenborn, A.M. (1983) *J.Magn.Reson. 53*, 423-442.

Clore, G.M. & Gronenborn, A.M. (1991a) *Prog.NMR Spec. 23*, 43-92.

Clore, G.M. & Gronenborn, A.M. (1991b) *Ann.Rev.Biophys.Biophys. Chem. 20*, 29-63.

Clore, G.M. & Gronenborn, A.M. (1993) *NMR of Proteins*. CRC Press, Boca Raton.

Cornell, B.A., Keniry, M., Hille, R.G. & Smith, R. (1980) *FEBS Letters 115(1)*, 134-138.

Cotrait, P.M. (1971) *Acta Cryst. B28*, 781-785.

Cung, M.T., Demange, P., Marraud, M., Tsikaris, V., Sakarellos, C., Papadouli, I., Kokla, A. & Tzartos, S.J. (1991) *Biopolymers 31*, 769-776.

Dauber-Osguthorpe, P., Roberts, V.A., Osguthorpe, K.J., Wolff, J., Genest, M. & Hagler, A.T. (1988) *Proteins: Structure Function Genetics 4*, 31-47.

Davis, D.G., Perlman, M.E. & London, R.E. (1994) (UnPub)

De Groot, H.J.M., Smith, S.O., Kolbert, A.C., Courtin, J.M.L., Winkel, C., Lugtenburg, J., Herzfeld, J. & Griffin, R.G. (1991) *J. Magn.Reson. 91*, 30-38.

De Vos, A.M., Ultsch, M. & Kossiakoff, A.A. (1992) *Science 255*, 306-312.

Dinauer, M.C., Orkin, S.H., Brown, R., Jesaitis, A.J. & Parkos, C.A. (1987) *Nature 327*, 717-720.

Dinauer, M.C., Pierce, E.A., Bruns, G.A.P., Curnutte, J.T. & Orkin, S.H. (1990) *J.Clin.Invest. 86*, 7303-7309.

Dinauer, M.C. (1993) *Crit.Rev.Clin.Sci. 30*, 329-369.

Dratz, E.A., Furstenau, J.E., Lambert, C.G., Thireault, D.L., Rarick, H., Schepers, S., Pakhlevaniants, S. & Hamm, H.E.

(1993) *Nature 363*, 276-280.

Dubois, B.W. & Evers, A.S. (1992) *Biochem. 31*, 7069-7076.

Dyson, H.J. & Wright, P.E. (1991) *Ann.Rev.Biophys.Biophys.Chem. 209*, 519-538.

Engelhard, M., Hess, B., Emeis, D., Metz, G., Kreutz, W. & Siebert, F. (1989a) *Biochem. 28*, 3967-3975.

Engelhard, M., Pevec, B. & Hess, B. (1989b) *Biochem. 28*, 5432-5438.

Englander, S.W. & Mayne, L. (1992) *Ann.Rev.Biophys.Biomol.Struct. 21*, 243-265.

Fahmy, K., Siebert, F. & Tavan, P. (1991) *Biophys.J. 60*, 989-1001.

Farrar, T.C. & Becker, E.D. (1971) *Pulse and Fourier Transform NMR*. Academic Press, New York.

Fesik, S.W. & Zuiderweg, E.R.P. (1990) *Quar.Rev.Biophys. 23(2)*, 97-131.

Fyfe, C.A. (1983) *Solid State NMR for Chemists*. CFC Press, Guelph Ontario.

Ganapathy, S., Chacko, V.P., Bryant, R.G. & Etter, M.C. (1986) *J. A.C.S 108(12)*, 3159-3165.

Gariepy, J., Kay, L.E., Kuntz, I.D., Sykes, B.D. & Hodges, R.S. (1985) *Biochem. 24*, 544-550.

Gerig, J.T. & Stock, A.D. (1975) *Org.Magn.Reson. 7*, 249-255.

Glaudemans, C.P.J., Lerner, L., Daves, G.D., Kovac, P., Venable, R. & Bax, A. (1990) *Biochem. 29*, 10906-10911.

Heinrikson, R.L. & Meredeth, S.C. (1984) *Anal.Biochem. 136*, 65-74.

Helgerson, S.L. & Dratz, E.A. (1988) *Biophys.J. 50*, 439a.

Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E. & Downing, K.H. (1990) *J.Mol.Biol. 213*, 899-929.

Herzberg, O. & James, M.N.G. (1985) *Biochem. 24*, 5298-5302.

Herzfeld, J. & Berger, A.E. (1980) *J.Chem.Phys. 73(12)*, 6021-6030.

Heyworth, P.G., Curnette, J.T., Nauseef, W.M., Volpp, B.D., Pearson, D.W., Rosen, H. & Clark, R.A. (1991) *J.Clin.Invest. 87*, 352-356.

Ikura, M., Spera, S., Barbato, G., Kay, L.E., Krinks, M. & Bax, A. (1991) *Biochem. 30*, 9216-9228.

Ikura, M., Clore, G.M., Gronenborn, A.M., Zhu, G., Klee, C.B. & Bax, A. (1992) *Science 256*, 632-637.

Jagannathan, N.R. (1989) *Magn.Reson.Chem. 27*, 941-946.

James, T.L. (1976) *Biochem. 15*, 4724-4730.

Jardetzky, O. & Roberts, G.C.K. (1981) *NMR in Molecular Biology.* Academic Press, New York.

Jeener, J., Meier, P., Bachmann, P. & Ernst, R.R. (1979) *J.Chem. Phys. 11*, 4546-4553.

Jonas, R. & Ebrey, T.G. (1991) *Proc.Natl.Acad.Sci. 88*, 149-153.

Kalk, A. & Berendenson, H.J.C. (1976) *J.Magn.Reson. 24*, 343-366.

Katre, N.V., Kimura, Y. & Stroud, R.M. (1986) *Biophys.J. 50*, 277-284.

Keepers, J. & James, T.L. (1984) *J.Magn.Reson. 57*, 404-426.

Kessler, H., Mronga, S., Muller, G., Moroder, L. & Huber, R. (1991) *Biopolymers 31*, 1189-1204.

Kleinberg, M.E., Malech, H.L. & Rotrosen, D. (1990) *J.Biol.Chem. 265*, 15577-15583.

Kleinberg, M.E., Mital, D., Rotrosen, D. & Malech, H.L. (1992) *Biochem. 31*, 2686-2690.

Kline, A., Braun, W. & Wuthrich, K. (1986) *J.Mol.Biol. 189*, 377-382.

Knaus, U.G., Heyworth, P.G., Evans, T., Curnutte, J.T. & Bokoch, G.M. (1991) *Science 254*, 1512-1515.

Koetzle, T.F. & Hamilton, W.C. (1972) *Abstract from Albuquerque, N. M. ACA abstract 07.*, 489-503.

Koshkin, V. & Pick, E. (1994) *FEBS Letters 338*, 285-289.

Kupke, D.W. & Fox, J.W. (1989) *Biochem. 28*, 4409-4415.

Kupke, D.W. & Shank, B.S. (1989) *J.Phys.Chem. 93*, 2101-2106.

Landy, S.B. & Rao, B.D.N. (1989) *J.Magn.Reson. 81*, 371-377.

Leto, T.L., Garrett, M.C., Fujii, H. & Nunoi, H. (1991) *J.Biol. Chem. 266*, 19812-19818.

Lind, M.D., Byungkook, L. & Hoard, J.L. (1965) *J.A.C.S 198(12)*, 1611-1612.

Lipari, G. & Szabo, A. (1982) *J.A.C.S 104* 4546-4559.
Lippens, G. & Hallenga, J.K. (1990) *J.Magn.Reson. 88*, 619.

Lippens, G.M., Cerf, C. & Hallenga, K. (1992) *J.Magn.Reson. 99*, 268-281.

Lippens, G.M., Hallenga, K., Belle, D.V., Wodak, S.J., Nirmala, N.R., Hill, P., Russel, K.C. & Hruby, V.J. (1993) *Biochem. 32*, 9423-9434.

Macura, S., Huang, Y., Suter, D. & Ernst, R.R. (1981) *J.Magn. Reson. 43*, 259-281.

Marion, D. & Bax, A. (1988) *J.Magn.Reson. 79*, 352-356.

Marion, D. & Wuthrich, K. (1983) *Biochem.Biophys.Res.Commun. 113*, 967-974.

Marsden, B.J., Hodges, R.S. & Sykes, B.D. (1988) *Biochem. 27*, 4198-4206.

Marsden, B.J., Hodges, R.S. & Sykes, B.D. (1989) *Biochem. 28*, 8839-8847.

Martin, M.L., Delpuech, J.J. & Martin, G.J. (1980) *Practical NMR Spectroscopy*. Heyden & Son Ltd, London.

McConnell, A.A., Nuttall, R.H. & Stalker, D.M. (1978) *Talanta 25*, 425-434.

Mehring, M. (1983) *High Resolution NMR in Solids*. Springer-Verlag Berlin Heidelberg, Berlin.

Meyer, E.F., Clore, G.M., Gronenborn, A.M. & Hansen, H.A. (1988) *Biochem. 27*, 725-730.

Milon, A., Miyazawa, T. & Higashijima, T. (1990) *Biochem. 29*, 65-75.

Moody, M.F., Jones, P.T., Carver, J.A., Boyd, J. & Campbell, I.D. (1987) *J.Mol.Biol. 193*, 759-774.

Morel, F., Doussiere, J. & Vignais, P.V. (1991) *Eur.J.Biochem. 201*, 523-546.

Murali, N., Jarori, G.K., Landy, S.B. & Rao, B.D. (1993) *Biochem. 32*, 12941-12948.

Nakanishi, A., Imajohohmi, S., Fujinawa, T., Kikuchi, H. & Kanegasaki, S. (1992) *J.Biol.Chem. 267*, 19072-19074.

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

Ni, F., Konishi, Y. & Scheraga, H.A. (1990) *Biochem. 29*, 4479-4489.

Ni, F., Ripoll, D.R., Martin, P.D. & Edwards, B.F.P. (1992) *Biochem. 31*, 11551-11557.

Nicholson, L.K., Teng, Q. & Cross, T.A. (1991) *J.Mol.Biol. 218*, 621-636.

Nilges, M., Clore, G.M. & Gronenborn, A.M. (1988) *FEBS Letters 239(1)*, 129-136.

Nirmala, N.R., Lippens, G.M. & Hallenga, K. (1992) *J.Magn.Reson. 100*, 25-42.

Nunoi, H., Rotrosen, D., Gallin, J.I. & Malech, H.L. (1988) *Science 242*, 1298-1301.

Nuttall, R.H. & Stalkir, D.M. (1977) *Talanta 24*, 355-360.

Opella, S.J., Stewart, P.L. & Valentine, K.G. (1987) *Quar.Rev. Biophys. 19*, 7-49.

Osapay, K. & Case, D.A. (1991) *J.A.C.S. 113*, 9436-9444.

Parkos, C.A., Allen, R.A., Cochrane, C.G. & Jesaitis, A.J. (1987) *J.Clin.Invest. 80*, 732-742.

Parkos, C.A., Allen, R.A., Cochrane, C.G. & Jesaitis, A.J. (1988) *Biochim.Biophys.Acta. 932*, 71-83.

Parkos, C.A., Dinauer, M., Jesaitis, A.J., Orkin, S. & Curnutte, J.T. (1989) *Blood 73*, 1416-1420.

Pervushin, K.V., Arseniev, A.S., Kozhich, A.T. & Ivanov, V.T. (1991) *J.Biomol.NMR 1*, 313-322.

Plateau, P. & Gueron, M. (1982) *J.A.C.S 104*, 7310-7311.

Powers, R., Garrett, D.S., March, C.J., Frieden, E.A., Gronenborn, A.M. & Clore, G.M. (1992) *Science 256*, 1673-1676.

Quinn, M.T., Parkos, C.A., Walker, L., Orkin, S.H., Dinauer, M.C. & Jesaitis, A.J. (1989a) *Nature 342*, 198-200.

Quinn, M.T., Parkos, C.A. & Jesaitis, A.J. (1989b) *Biochim. Biophys. Acta. 987*, 83-94.

Quinn, M.T., Curnutte, J.T., Parkos, C.A., Mullen, M.L., Scott, P.J., Erickson, R.W. & Jesaitis, A.J. (1992a) *Blood 79*, 2438-2445.

Quinn, M.T., Mullen, M.L. & Jesaitis, A.J. (1992b) *J.Biol.Chem. 267*, 7303-7309.

Quinn, M.T., Evans, T., Loetterle, L.R., Jesaitis, A.J. & Bokoch, G.M. (1993) *J.Biol.Chem. 268*, 20983-20987.

Raleigh, D.P., Grey, C.P., Soffe, N. & Dobson, C.M. (1992) *J. Magn.Reson. 97*, 162-170.

Redfield, A.G. & Kuntz, S.D. (1975) *J.Magn.Reson. 19*, 250-254.

Rees, D.C., Komiya, H., Yeates, T.O., Allen, J.P. & Fehr, G. (1989) *Annu.Rev.Biochem. 58*, 607-633.

Reid, R.E., Gariepy, J., Saund, A.K. & Hodges, R.S. (1981) *J. Biol.Chem. 256(6)*, 2742-2751.

Reid, R.E. (1990) *J.Biol.Chem. 265(11)*, 5971-5976.

Rhodes, G. (1993) *Crystallography Made Crystal Clear*. Academic Press, San Diego CA.

Richardson, J.S. (1981) *Advances in Protein Chemistry 34*, 167-300.

Rotrosen, D., Yeung, C.L., Leto, T.L., Malech, H.L. & Kwong, C.H. (1992) *Science 256*, 1459-1462.

Scherf, T., Hiller, R., Naider, F., Levitt, M. & Anglister, J. (1992) *Biochem. 31*, 6884-6891.

Scott, W.G., Milligan, D.L., Milburn, M.V., Prive, G.G., Yeh, J., Koshland, D.E. & Kim, S. (1993) *J.Mol.Biol. 232(2)*, 555-573.

Segal, A.W. (1987) *Nature 326*, 88-91.

Segal, A.W. (1989) *J.Clin.Invest. 83*, 1785-1793.

Segal, A.W., West, I., Wientjes, F., Nugent, J.H.A., Chavan, A.J., Haley, B., Garcia, R.C., Rosen, H. & Scarce, G. (1992) *Biochem. 284*, 781-788.

Sheldrick, G.M. (1983) (UnPub)

Sievers, R. (1973) *Nuclear Magnetic Resonance Shift Reagents*. Academic Press, New York.

Smith, R.L. & Oldfield, E. (1984) *Science 225*, 280-288.

Smith, R.M. & Curnutte, J.T. (1991) *Blood 77*, 673-686.

Smith, S.O. & Peersen, O.B. (1992) *Annu.Rev.Biomol. Struct. 21*, 25-47.

Solomon, I. (1955) *Phy. Rev. 99*, 559.

Stilbs, P. & Moseley, M.E. (1978) *J.Magn.Reson. 31*, 55-61.

Strynadka, N.C.J. & James, M.N.G. (1989) *Annu.Rev.Biochem. 58*, 951-998.

Sukumar, M. & Higushijima, T. (1992) *J.Biol.Chem. 267(3)*, 21421-21424.

Teahan, C., Rowe, P., Parker, P., Totty, N. & Segal, A.W. (1987) *Nature 327*, 720-721.

Tsang, P., Rance, M., Fieser, T.M., Ostresh, J.M., Houghten, R.A., Lerner, R.A. & Wright, P.E. (1992) *Biochem. 31*, 3862-3871.

Tycko, R. & Dabbagh, G. (1991) *J.A.C.S. 113*, 9444-9448.

Vajda, S., Jafri, M.S., Sezerman, O.U. & DeLisi, C. (1993) *Biopolymers 33*, 173-192.

Vanberkum, M.F.A. & Means, A.R. (1991) *J.Biol.Chem. 266*, 21488-21495.

Veeman, W.S. (1984) *Prog.in NMR Spec. 16*, 193-235.

Volpp, B.D., Nauseef, W.M. & Clark, R.A. (1988) *Science 242*, 1295-1297.

Volpp, B.D., Nauseef, W.M., Donelson, J.E., Moser, D.R. & Clark, R.A. (1989) *Proc.Natl.Acad.Sci. 86*, 7195-7199.

Vyas, N.K., Vyas, M.N. & Quiocho, F.A. (1987) *Nature 327(18)*, 635-638.

Wagner, G., Hyberts, S.G. & Havel, T.F. (1992) *Ann.Rev.Biomol. Struct. 21*, 167-198.

Wang, Z., Jones, J.D., Rizo, J. & Gierasch, L.M. (1993) *Biochem. 32*, 13991-13999.

Weast, R.C. (1986) *CRC Handbook of Chemistry and Physics*. CRC Press, Boca Raton.

Weiner, P.K. & Kosslman, P.A. (1981) *J.Computation.Chem. 2*, 287.

Weiss, M.S., Abele, U., Weckesser, J., Welte, W., Schiltz, E. & Schulz, G.E. (1991) *Nature 307*, 1627-1629.

Wenzel, T.J. (1987) *NMR Shift Reagents*. CRC Press, Inc., Boca Raton.

Williams, J.S. & Rosevear, P.R. (1991) *J.Biol.Chem. 266(4)*, 2089-2098.

Wilmot, C.M. & Thornton, J.M. (1988) *J.Mol.Biol. 203*, 221-232.

Wishart, D.S., Sykes, B.D. & Richards, F.M. (1991) *J.Mol.Biol. 222*, 311-333.

Withka, J.M., Srinivasan, J. & Bolton, P.H. (1992) *J.Magn.Reson. 98*, 611-617.

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

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

Yan, B., Nakanishi, K. & Spudich, J.L. (1991a) *Proc.Natl.Acad. Sci. USA 88*, 9412-9416.

Yan, M., Manor, D., Weng, G., Chao, H., Rothberg, L., Jedju, T.M., Alfano, R.R. & Callender, R.H. (1991) *Proc.Natl.Acad.Sci. USA 88*, 9809-9812.

Zheng, J. & Post, C.B. (1993) *J.Magn.Reson. Series b 101*, 262-270.

HOUCHEN
BINDERY LTD