To The Graduate Council:

I am submitting herewith a dissertation written by John Allan Lamerdin entitled "Site-Specific Isotope Labeling Studies of Yeast Phosphoglycerate Kinase." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biochemistry.

Engin Serpersu

Dr. Engin Serpersu, Major Professor

We have read this dissertation
and recommend its acceptance.

Elizabeth E. Howell

Cynthia B. Peterson

Scott K. Niyogi

Accepted for the Council:

Associate Vice Chancellor and
Dean of The Graduate School

SITE-SPECIFIC ISOTOPE LABELING STUDIES OF YEAST PHOSPHOGLYCERATE KINASE

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

John Allan Lamerdin

December, 1997

DEDICATION

This dissertation is dedicated to my parents

Richard and Isabel Lamerdin

with appreciation for their love, support and encouragement.

ACKNOWLEDGMENTS

First and foremost I would like to express my deep appreciation to my mentor, Dr. Engin Serpersu. I thank him for his patience, guidance and encouragement. I also thank him for providing me with an opportunity to learn and to grow in many different ways. I thank Dr. Lee Magid for her insightful comments and suggestions on my work. I also thank Dr. Salil Niyogi for a productive and interesting collaboration on human epidermal growth factor. Dr. Stephen Henderson also deserves special thanks for contributing his expertise on the SANS work. I would also like to thank Dr. Cynthia Peterson and Dr. Liz Howell for their helpful input and direction of my research, as well as for their friendship, support and encouragement. I hope they know that both were equally invaluable.

I must also thank my friends, lab mates and professors who helped me keep going through the difficult times, were always ready with a smile or a joke, and made (almost) every day enjoyable: Tonia Lane, Dr. Kamesh Pappu, Enrico DiGiammarino, Ayca Akal, Christine Schar, Dr. Baburaj Kunnumal, Angelia Gibson, Dr. Amanda Schultz, Robby Brooks, Laurel O'Connor, Dr. Barry Bruce and Dr. Dan Roberts.

I would also like to thank my family: my parents, Richard and Isabel Lamerdin, my brother, Stewart Lamerdin, and especially my sister, Jennifer Lamerdin. When things looked the bleakest, they were always there to help me back up. Without their love, encouragement and prayers, I would not have completed my studies. I hope to someday give back to them what they have given to me over these last five years.

Abstract

Yeast phosphoglycerate kinase (PGK) is a 47 kDa glycolytic enzyme composed of 415 amino acids. By nuclear magnetic resonance (NMR) spectroscopic standards, this is a large protein, and difficult to study using conventional NMR techniques. To demonstrate the feasibility of the study of large proteins by NMR, three isotopically different forms of yeast PGK were biosynthetically prepared. Firstly, perdeuterated yeast PGK was prepared that contained isotopically normal tyrosine residues. This enzyme was characterized using one and two dimensional NMR experiments. Titration studies of this enzyme with the substrates MgATP and the competitive inhibitor glycerol-3-phosphate (G3P) indicated that at least one of the labeled tyrosine residues was located in close spatial proximity to the MgATP binding site, and whose local magnetic environment was altered upon substrate binding. Addition of G3P had minimal effect on the conformation of the enzyme.

The second enzyme to be biosynthetically prepared was perdeuterated PGK containing $^{13}C_{\alpha}$ tyrosine, ^{15}N phenylalanine and ^{15}N isoleucine. This enzyme was prepared in order to permit identification of two of the seven tyrosine residues found in the primary sequence of PGK. Initial characterization of this enzyme by NMR techniques was successful, however conclusive identification of the tyrosine residues was not achieved.

The third enzyme to be biosynthetically prepared was perdeuterated PGK containing $^{13}C_{\alpha}$ tyrosine and ^{15}N valine. This enzyme was also prepared to permit

identification of two tyrosine residues. Using one, two and three dimensional NMR techniques, tyrosine residues 48 and 56 were conclusively identified in NMR spectra of this enzyme. This work represents the first successful attempt to conclusively identify two individual residues of a large protein using NMR methodology. It further demonstrates the feasibility structural studies of large proteins using nontraditional isotope labeling strategies in conjunction with traditional NMR techniques.

Expression of yeast PGK in an auxotrophic system would greatly enhance the incorporation levels of isotopes included in the growth media of the cells. However, the availability of auxotrophic yeast strains is limited. It was decided, therefore, to express the yeast gene in a bacterial system. The yeast PGK gene was cloned and ligated into a vector containing the T7 RNA polymerase promoter site. *E. coli* cells of the strain BL21(DE3)pLysS were then transformed with this construct. It was observed that, although the yeast protein was expressed, the expression levels obtained were comparable to those achievable using the yeast system.

The association of yeast PGK with the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was also investigated using small-angle neutron scattering (SANS) techniques. It has been proposed that these enzymes, which are successive in the glycolytic pathway, form a complex in the cell in order to facilitate transfer of the PGK reaction product 1,3-bisphosphoglycerate. Preliminary analysis of SANS data suggests that these enzymes do associate to allow such a transfer.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADP	Adenosine 5'-diphosphate
AMP-PNP	Adenylimidodiphosphate
ATP	Adenosine 5'-triphosphate
COSY	Correlated spectroscopy
dATP	2'-Deoxyadenosine 5'-triphosphate
DEPT	Distortionless enhancement by polarization transfer
DNAse	Deoxyribonuclease
DSS	Sodium 2,2-dimethyl-2-silapentane-5-sulfonic acid
G3P	Glycerol-3-phosphate
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
HNCA	Triple resonance ^{1}H-^{15}N-$^{13}C_{\alpha}$ correlation experiment
HSQC	Heteronuclear single quantum coherence
Ile	Isoleucine
INEPT	Insensitive nuclei enhanced by polarization transfer
INV4TP	Inverse-detected $^{13}C_{\alpha}$-H_{α} correlation experiment
INV4MLNDTP	Inverse-detected, MLEV-17-based $^{13}C_{\alpha}$-H_{α}-H_{β} correlation experiment
Mg	Magnesium
NAD^{+}	Nicotinamide adenine dinucleotide

NADH	Nicotinamide adenine dinucleotide, reduced form
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhaüser effect
NOESY	Nuclear Overhaüser effect spectroscopy
3PGA	3-Phosphoglycerate
PGK	Phosphoglycerate kinase
Phe	Phenylalanine
R_G	Radius of gyration
RNAse	Ribonuclease
SANS	Small-angle neutron scattering
S_N2	Bimolecular nucleophilic substitution
TOCSY	Total correlation spectroscopy
Tris	Tris (hydroxymethyl) aminoethane
Tyr	Tyrosine
Val	Valine

Chapter I

Introduction

Yeast phosphoglycerate kinase (EC 2.7.2.3) is a 47 kDa monomeric glycolytic enzyme that catalyzes the reversible transfer of a phosphate group from ATP to 3-phosphoglycerate, giving ADP and 1,3-bisphosphoglycerate:

$$\text{3-phosphoglycerate} + \text{ATP} \overset{Mg^{2+}}{\Leftrightarrow} \text{1,3-bisphosphoglycerate} + \text{ADP}$$

This reaction requires the presence of magnesium, and is the first step in the glycolytic pathway to generate ATP.

Structural analyses of yeast PGK (Watson *et al.*, 1982, Bryant *et al.*, 1974) show the protein to be composed of two lobes connected by a short sequence. The yeast enzyme is composed of 14 alpha helices and 14 beta sheets, and shares a high degree of tertiary homology with PGK from pig muscle (Harlos *et al.*, 1992), *Bacillus stearothermophillus* (Davies *et al.*, 1994) and horse muscle (Blake & Rice, 1981). In yeast PGK, each domain is composed of a central core of six parallel strands of β-sheet, which are surrounded by helices and are coupled to the sheets by means of β turns and sequences of random coil (Figure I-1)*.

The catalytic mechanism of PGK has been intensely studied, especially in light of

---

* All figures may be found in Appendix A and all tables may be found in Appendix B.

its unique bilobal structure. The exact substrate binding sites have not as yet been precisely identified. The 3-PGA binding site has been localized to the positively charged basic patch in the pig muscle enzyme (Harlos *et al.*, 1992) on the N terminal lobe of the enzyme. Previous studies of the yeast enzyme, however, do not agree with this data (Watson *et al.*, 1982).

The MgATP binding site has been proposed to be on the C terminal lobe of the enzyme, facing the PGA binding site (Watson *et al.*, 1992). This study purported that the metal binding site is likely to involve Asp 372, although recent data (Pappu *et al.*, 1997) indicate that the metal is probably coordinated to Glu 398. The proposed MgATP binding site consists of 6 beta sheets, which is consistent with the nucleotide binding sites of other kinases.

The catalytic mechanism of PGK has been determined to be a direct transfer mechanism (Webb & Trentham, 1980). This study utilized ATP whose γ phosphate group was labeled with O^{16}, O^{17} and O^{18}. The results of this study show that catalysis proceeds with an inversion of configuration, a hallmark of direct transfer mechanisms. The possibility of an S_N2-type mechanism is also suggested.

The substrate binding sites on PGK have been determined, by NMR, analysis of the crystal structure, and other techniques, to be distant from each other: the γ phosphate of ATP and PGA are separated by a distance of about 10 Å. This distance is clearly too great to permit the direct transfer of the phosphate group. In order to correlate the observed data, a "hinge-bending" mechanism was proposed (Banks *et al.*, 1979). In this

scheme, upon binding substrates, each lobe of the enzyme flexes about the short connector region to bring the substrates in close proximity to each other. The accuracy of this scheme to explain the catalysis of PGK has been a subject of great debate (e.g. McPhillips *et al.*, 1996). McPhillips *et al.* argue, on the basis of data obtained from the crystal structure of a mutant PGK protein crystallized in a ternary form, that the hinge-bending mechanism inadequately explains the observed catalysis. It should be noted that, in this study, the enzyme was cocrystallized with the substrate analog Mg-AMP-PNP, and not MgATP. Also, this analog was modeled into the structure and its conformation was not determined from the X-ray crystallography data.

Several studies have indicated that substrate binding is accompanied by a substantial conformational change in the structure of PGK. Pickover *et al.* (1979) used small-angle X-ray scattering to demonstrate that the radius of gyration of the ternary complex of PGK, MgATP and PGA (R_G = 23.3 Å) decreases by 0.7 Å compared to the radius of gyration of the enzyme in the absence of substrates (R_G =24.0 Å). Similar results were obtained from small-angle neutron scattering experiments (Henderson *et al.*, 1994). Henderson *et al.* found that the radius of gyration of the ternary complex was decreased 1.2 ± 0.3 Å compared to the radius of gyration of the enzyme alone. Sedimentation experiments (Roustan *et al.*, 1980) also show that the enzyme becomes more compact in the presence of substrates. Dryden *et al.* (1992) used fluorescence methodology to arrive at the same conclusion. NMR (Tanswell *et al.*, 1976) has also been employed to study, and confirm, this conformational change.

The fluorescence study of Dryden *et al.* (1992) provided insight into the specific steps of the hinge-bending mechanism that take the enzyme from its open form, in which no substrates are bound, to its closed form, when the ternary complex is formed. Dryden *et al.* suggest that the hinge region is inflexible at neutral pH, and becomes more labile as substrates bind. In this study, the fluorescent probe was observed to reorient by 1-5° in the presence of substrates. This change was interpreted as an increase in the hinge region upon binding substrates. This increase in flexibility allows the two domains to come together and catalysis to proceed. This data served to refine the permissive model and refuted earlier hypotheses that substrate binding directed specific motions of the hinge (Adams *et al.*, 1985).

The invocation of a hinge-bending mechanism is not limited to PGK—in fact it has been found to be a common mechanism for many other kinases. Bennett and Steitz (1978) found that upon binding glucose, yeast hexokinase, which has a bilobal structure similar to PGK, exhibits a 12 degree rotation of one of its lobes. In the case of yeast PGK, a rotation of similar magnitude would explain the observed decrease in the radius of gyration seen by Pickover *et al.* (1979) in their X-ray scattering studies of the yeast PGK ternary complex.

The crystal structure of the ternary closed form of the yeast enzyme has not yet been achieved. Recently, the structure of yeast PGK in the presence of 3-PGA and the substrate analog Mg-AMP-PNP has been determined (McPhillips *et al.* 1996). This study provides insight into the binding site of 3-PGA, however it does not represent the closed structure of the enzyme, since the true substrates are not present. Moreover, previous

experiments in the Serpersu laboratory suggest that AMP-PNP is not an accurate analog of ATP. Similarly, the crystal structure of the pig muscle enzyme (Harlos *et al.*, 1992) was crystallized in the presence of 3-PGA, but no nucleotide was present. Again, this structure represents the open form or, at best, a partially closed form of the enzyme.

Early in 1997, Bernstein *et al.* successfully crystallized and solved the structure of the ternary complex (3-PGA-Mg-ADP) of PGK isolated from *Trypanosoma brucei.* Their results show that PGK does, in fact, undergo a hinge-bending mechanism, bringing the substrates in close proximity to each other. They found that binding of 3-PGA alone induces an eight degree bend in the hinge. When all substrates are bound, there is an overall bending of 32 degrees and an overall displacement of 27 Å of the ends of each domain. By combining the structural data from their work with data obtained from other studies (Blake & Evans, (1974), Harlos *et al.*, (1982), Davies *et al.*, (1994)), Bernstein *et al.* propose a model describing the conformational changes undergone by PGK upon binding substrates (Figure I-2). This model suggests that the most significant conformational changes occur in the hinge region of the enzyme, and this region moderates closure of the enzyme.

A large volume of information on the closed form of the yeast enzyme has come from the Serpersu laboratory. In 1993, Gregory and Serpersu determined the active site geometry of the ternary complex. They found that the distance between 3-PGA and the paramagnetic substrate analog CrATP was consistent with the hinge-bending mechanism, and that the substrates were in close enough proximity to allow a direct transfer of the phosphate group. Pappu *et al.* (1994) further investigated the closed conformation of

yeast PGK and found that both the 3-PGA and ATP were coordinated to the metal ion in the ternary complex. Both of these studies utilized NMR spectroscopy as a primary technique, demonstrating the utility of this methodology for the study of the closed conformation of PGK.

NMR spectroscopy is one of the most powerful methods available for obtaining structural information about proteins. However, it is limited by the size of the protein being studied. Large proteins, from an NMR standpoint, are virtually impossible to study with traditional NMR techniques, owing to the complex spectra they generate. In order to facilitate further structural studies of yeast PGK, the work described in this dissertation was undertaken. The main goal of this research is to provide a methodology to allow the study of large proteins by means of NMR spectroscopy. Two different labeling schemes are presented. The first has the ultimate goal of providing insight into the interactions of PGK with its substrates in the ternary complex. The second represents further refinement of the methodology and is the first demonstration of a labeling strategy to allow site specific identification of two unique residues out of the 415 amino acid sequence of PGK. This latter strategy was developed to allow detailed structural studies of proteins that are too large to be studied using conventional NMR methods. An additional approach, related to this research and is presented in this document, is an attempt to express yeast PGK in an *Escherichia coli* expression system. Finally, a study of the interaction of yeast PGK with the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase, using small-angle neutron scattering approaches, is also presented.

Chapter II

Studies of Protonated Tyrosine-Labeled Perdeuterated PGK

A. Introduction

NMR spectroscopy is one of the most powerful techniques available for the study of proteins and their interactions in solution. NMR spectroscopy can provide detailed information on a protein's structure in solution and can allow an assessment of the effect of ligands on the structure of a protein. NMR spectroscopy also allows insight into the kinetic and thermodynamic properties of a protein or peptide. Another benefit of NMR is that it provides a dynamic view of the structure of a protein, rather than the static picture obtainable from X-ray crystallographic studies. Indeed, there have been a number of cases in which the structural information obtained from crystallographic studies has been revised by NMR-based experiments. Another advantage NMR spectroscopy has over X-ray crystallographic studies is that the data obtained from an NMR study are immediately accessible; X-ray crystallographic studies require successful crystallization of the sample, preparation of appropriate heavy-atom derivatives and the phasing problem must be solved.

NMR studies typically involve initial characterization of the one-dimensional spectrum of a protein. This is followed by two-dimensional experiments including COSY experiments, which rely on through-bond connectivities, and NOESY experiments, which rely on the through-space transfer of magnetization from one

magnetic nucleus to another. NOESY experiments are of particular interest in structural studies because the intensity of the observed crosspeak is a function of the distance between the two nuclei, with crosspeak intensity related to distance by a factor of $1/r^6$, where r is the distance separating the two nuclei. If the distance between two magnetic nuclei is known and the crosspeak corresponding to that interaction can be identified in the spectrum, that known distance can then be used as a calibration or "spectral ruler" for the determination of distances between other nuclei. What emerges from this type of analysis is a three dimensional picture of exactly how each nucleus is positioned relative to the other nuclei. If the resonances of ligands associated with the sample can be identified, the orientation of the ligand with respect to the other nuclei can also be established.

There are two main drawbacks to NMR spectroscopy: one is the amount of sample required to provide useful spectra, and the other is the limitation on the size of the protein that can be studied. Presently, the upper size limit of proteins having structures that can be solved using conventional NMR techniques is about 25 kDa. There are a number of complications associated with the analysis of proteins with molecular weights larger than this. Firstly, as the size of the protein increases, the number of resonances increases as well, and leads to complicated, unresolved spectra. Secondly, the presence of a large number of protons and the longer tumbling times associated with these larger proteins can lead to broadening of the peaks of the spectra, and the loss of spectral information. Thirdly, spin diffusion becomes a very real problem for large proteins: During a NOESY experiment, magnetization is transferred from one

nucleus to another, and the buildup of magnetization on the second nucleus is proportional to $1/r^6$, where r is the distance between the two nuclei. This known buildup rate enables structural analysis of a system based on NOE intensities. However, it is possible for the magnetization that has built up on the second nucleus to be transferred to a third nucleus (LeMaster, 1990). This second transfer of magnetization is referred to as "spin diffusion". Spin diffusion results in the appearance of crosspeaks in the NOESY spectrum that that are indistinguishable from primary NOE crosspeaks and can greatly complicate a distance analysis. As the mixing time in a NOESY experiment is increased, the amount of spin diffusion that occurs also increases. In order to limit the degree of spin diffusion, it is necessary to shorten the mixing time of the experiment, and thus, sacrifice the sensitivity of the experiment. Obviously, for a protein such as PGK, which has about 4000 protons, spin diffusion becomes a very serious concern for NMR studies, and can significantly complicate or prevent interpretation of one and two dimensional spectra.

Figure II-1 is a proton NMR spectrum of yeast PGK. This spectrum represents all 4000 proton resonances of the 415 residues of this protein. Clearly, the low resolution and overlap of the resonances preclude any kind of detailed analysis of this spectrum. Furthermore, any study of the effect of substrate on this protein is impossible, due to the inability to distinguish substrate resonances from the protein background, as well as the inability to identify specific changes in the spectrum. The effect of spin diffusion on the results of NOESY-type experiments has been addressed above.

The most significant advance in the study of proteins by NMR is the ability to replace protons with deuterons. Deuterons resonate far away from the region of the spectrum where protons are observed and therefore, deuteration of a protein can dramatically simplify the proton spectrum of a sample. The advantage deuteration provides is demonstrated clearly in Figure II-2. This figure shows the aromatic regions of both protonated and perdeuterated yeast PGK. This ability to reduce the complicated protonated spectrum to a completely clear background opens up many possibilities for studies of substrate conformation in the presence of the perdeuterated enzyme (Shibata *et al.*, 1995). In the study of Shibata *et al.*, the resonances of the substrates were the only observed signals in the spectrum. As a result of the use of perdeuterated yeast PGK in this study, insight into the conformation of Mg(II)dATP when bound to the enzyme was obtained, information not accessible using the protonated enzyme.

Deuteration also has the benefit of being able to reduce the number of available relaxation pathways and therefore the deuterated protein exhibits longer relaxation times. The dipolar interactions of protons are the chief relaxation mechanism and by reducing the number of protons available, the longitudinal relaxation times (T_1's) become longer. The result of lengthening the relaxation times manifests itself as sharper, narrower lines in the spectrum.

Deuteration of a protein can furthermore be of great advantage in the minimization of crosspeaks observed in two-dimensional NOESY spectra that arise due to spin diffusion. Recall that spin diffusion arises from the transfer of magnetization from one nucleus to another by means of a third. More explicitly, magnetization can be

transferred from one nucleus to another. The magnetization can then build up on the second nucleus and be transferred to a third. This transfer shows itself as a crosspeak between the first and third nucleus. These "false" secondary NOE crosspeaks can complicate distance analyses if they are undistinguished from primary crosspeaks. Deuteration greatly reduces the extent of spin diffusion by eliminating the number of protons that are in close enough spatial proximity to accept transferred magnetization, those protons having been replaced with deuterons. The dipolar interactions between protons and deuterons are significantly less efficient (only 1/16th as efficient) than those between protons (LeMaster, 1989). Thus, the deuteration of proteins can enhance the analysis of NOESY spectra.

Deuteration of proteins was first accomplished in green algae (Crespi *et al.*, 1960). In this study, algae were grown on a deuterated media, providing perdeuterated proteins, sugars and lipids. The next phase in the development of a deuteration strategy for the study of proteins came in 1968 when Markley *et al.* (1968) successfully prepared a protein that was selectively deuterated at the side chain positions. Markley *et al.* used this method to make resonance assignments for a protein whose spectrum was far too complex to analyze using traditional NMR methods.

Another approach that has been utilized to simplify NMR spectra is partial fractional deuteration (LeMaster, 1989). Partial fractional deuteration is the non-specific, random deuteration of a sample. This technique was used to improve the resolution of two-dimensional correlated spectroscopic studies, and allowed analysis of spectra not normally interpretable due to peak overlap.

The above mentioned approaches are not adequate for detailed studies of yeast PGK for a number of reasons. Firstly, although partial deuteration can remove a significant number of resonances from a spectrum, a large number of signals remain. For example, deuteration of the side chains of a large protein such as PGK would markedly decrease the number of signals in a proton spectrum, but would leave amine and alpha-proton regions composed of 415 signals in each region, a clearly unmanageable number of signals. Partial deuteration would also reduce the number of observable signals, but would still leave a spectrum too complex to effectively analyze. Moreover, the residual protons would still allow for many different relaxation pathways, decreasing spectral resolution.

Another approach that has been employed for the analysis of large proteins by NMR is isotope labeling. In this conventional strategy, the protein is prepared with either all or just the backbone carbons being replaced with ^{13}C. Typically, the backbone amine nitrogens are also replaced with ^{15}N. Replacement of the backbone nuclei with these isotopes allows correlation of their resonances and permits the backbone resonance assignments to be made (e.g. Chan & Markley, 1982). This type of labeling is referred to as "uniform labeling", as each carbon and nitrogen of the protein has been nonspecifically labeled.

Perdeuteration has been successfully employed for the study of large proteins (Shibata *et al.*, 1995, Kalbitzer *et al.*, 1985), although these studies did not directly address structural details. In fact, although perdeuteration as a method for simplification of spectra is unsurpassed, it is of limited use in structural studies. When a protein is

perdeuterated, each of its protons is replaced with a less-sensitive deuteron. This reduces the degree of spectral overlap and the signals of a substrate become clearly visible. However, any structural information about the protein itself must come from the position and line shape of peaks due to protons attached to the protein. In the perdeuterated enzyme these protons are absent.

To date, only Pappu *et al.* (1994) have been able to extrapolate the advantages of perdeuteration to achieve the goal of structural details from an NMR study of a large protein. In this study, perdeuterated PGK was biosynthetically prepared, however the protein was further labeled with histidines (PGK has eight histidines in its primary sequence) of normal isotopic composition. This uniquely labeled protein exhibited the advantages of both the perdeuterated protein and a labeled protein: the background protein resonances were removed, and yet the introduction of protonated histidines gave eight protonated residues to act as probes of the enzyme's conformation. This labeling strategy allowed a ^{1}H NMR study of substrate-ligand interactions that showed that in the presence of substrates, the histidine probes shifted position up to 6 Å. Similar strategies have been employed previously (Crespi *et al.*, 1968; Crespi *et al.*, 1970), but with limited success. These earlier studies showed that such a labeling procedure was possible, but the spectra obtained from this work were not suitable for a detailed analysis due to broad linewidths and overlapping resonances. The work of Pappu *et al.* (1994) represents the first successful application of this specific labeling strategy to a large protein that has led to meaningful structural data.

This isotope-labeling strategy was further extended in the work described in this dissertation. Instead of the introduction of protonated histidine residues, perdeuterated PGK was prepared that contained protonated tyrosine residues. Tyrosine was selected as the label of choice for a number of reasons. Firstly, there are only seven tyrosine residues in the primary sequence of PGK. These tyrosines are located at positions 48, 56, 74, 122, 158, 193 and 380. Only tryptophan occurs less frequently in the primary sequence of this protein (Watson *et al.*, 1982). This represents a reduction of almost 4000 protons to approximately 60 protons. The simplification achieved through the use of this strategy is obvious. Concurrent with the reduction in signals to a more manageable number comes the elimination of a significant number of potential relaxation pathways. The reduction in the number of relaxation pathways gives longer longitudinal relaxation times and leads to not only to narrower linewidths, but makes it possible to study enzyme-substrate complexes at much lower ratios of enzyme to substrate.

Secondly, the possibility of overlap of enzyme and substrate peaks becomes a moot point: peaks in the ^{1}H NMR spectrum that arise from ATP are far removed from those that correspond to the α, β and ring protons of tyrosine. Peaks arising from the presence of 3PGA (or the substrate analog glycerol-3-phosphate (G3P)) are also far removed from those of tyrosine. Moreover, since there are only seven tyrosine residues and each is located in a different local magnetic environment, it is possible that the peaks of the tyrosine residues will show a minimal degree of overlap among themselves. The ring protons of tyrosine are especially susceptible to the local environment and offer the best chance for resolution.

Thirdly, it has been observed that chemical modification of the tyrosine residues of yeast PGK led to inactivation of the enzyme (Markland *et al.*, 1975; Roustan *et al.*, 1976; Roustan *et al.*, 1979). All three of these studies suggest that there is at least one tyrosine residue in PGK that is essential for catalysis. Of these, the study of Roustan *et al.* (1979) is of particular interest. In this study the reagent 7-chloro-4-nitrobenzofuran was used. It was observed that as the pH was raised from 7.3, the pH at which the critical tyrosine was modified, to 9, the reagent was observed to shift from the modified tyrosine to a lysine, suggesting that a lysine is in close spatial proximity to the critical tyrosine, near the catalytic site. This observation confirms earlier work (Markland *et al.*, 1975). In the study of Markland *et al.*, lysine residues were covalently modified with three different reagents: *O*-methylisourea, 2-methoxy-5-nitropone and pyridoxal phosphate. Each modification was performed on PGK, and each modification indicated the presence of three essential lysine residues. Substrate protection studies suggest that the substrate analog MgITP protects up to three lysines from modification, and 3-PGA protects one.

Lastly, the tyrosine residues are distributed throughout the structure of the protein. That is, they are located on both lobes of the protein (See Figure II-3), making them good probes of overall structural changes.

The aromatic resonances of tyrosine appear in the NMR spectrum at 6.86 ppm and 7.14 ppm. In the case of PGK, these two peaks represent a total of 28 protons (two at the 2,6 position and two at the 3,5 position of each of the seven tyrosines). These peaks arise from the splitting of the 2,6 protons by the 3,5 protons and vice versa. This splitting results in the total area of the protons being divided into two peaks. This splitting

therefore reduces the intensity of each of the peaks. In an effort to further simplify the spectrum and enhance the signal intensity of these peaks, tyrosine was prepared that was deuterated at the 2,6 positions, leaving the 3,5 positions protonated. By incorporating this partially deuterated tyrosine into the enzyme, the tyrosine ring signals in the NMR spectrum are reduced in number from two to one, the 2,6 peak becoming silent due to the exchange of the protons with the deuterons. Moreover, the splitting of the 3,5 peak is effectively eliminated due to the lesser degree of interaction between the deuteron and the proton, as compared to proton-proton interactions. This results in an increase in the signal intensity of the peak due to the 3,5 ring protons.

The primary objectives from this section of work are twofold: Firstly, selectively deuterated tyrosine will be prepared. This isotopically-modified residue will then be included in the deuterated yeast growth medium. PGK will then be isolated and characterized using NMR methodology. Secondly, the effect of the addition of substrates on the NMR spectrum of the tyrosine-labeled enzyme will be investigated.

B. Experimental Procedures

1. Materials

Phosphoglycerate kinase was purified from *Saccharomyces cerevisiae* strain 20-B12. This strain carries the multicopy plasmid pCGY219. Both the plasmid and strain were a gift of Dr. G. Hitzeman, formerly of Genentech.

Deuterium oxide (99.6%), deuterium chloride and sodium deuteroxide were purchased from Isotec, Inc. Deuterated algal cells were purchased from Cambridge Isotopes.

Glycerol-3-phosphate, glyceraldehyde-3-phosphate dehydrogenase, 2'-dATP, tyrosine and all nutrients included in the growth media, except yeast nitrogen base without amino acids, were purchased from Sigma. YNB without amino acids was purchased from Difco. 2'-dATP was passed over a Chelex-100 column to remove trace metals prior to use in experiments. The concentration was then determined spectrophotometrically.

NMR supplies, including tubes, caps and plugs, were purchased from Wilmad Glass Co.

2. Preparation of 2,6 Deuterated Tyrosine

2,6 deuterated tyrosine was prepared using the procedure of Wishart *et al.* (1993). This procedure requires activation of the platinum catalyst which was accomplished by adding 2 g of sodium borohydride to a suspension of platinum oxide (0.5 g) in 200 mL of water that had been extensively degassed. This mixture was stirred for 30 minutes and the reaction vessel warmed to 70° C to hydrolyze excess sodium borohydride. The flask was cooled, the water decanted, and the precipitate washed five times with aliquots of water. The catalyst was then washed with 10 mL of deuterium oxide.

About 0.5 g of the catalyst was then added to 30 mL of deuterium oxide. Two grams of tyrosine were added followed by 2 mL of sodium deuteroxide to solublize the

tyrosine. The mixture was then refluxed under nitrogen for 48 hours. The product was filtered hot to remove the catalyst for future use, and the solution was neutralized with deuterium chloride. The solution was cooled overnight at 4° C and the crystals were then collected and washed with cold deuterium oxide. The material was lyophilized and stored for later use as an isotope label in subsequent isotope-labeled PGK preparations.

3. Hydrolysis of Deuterated Algal Cells

Deuterated algal cells were refluxed at 100° C in 6 N deuterium chloride. The hydrolysate was then cooled, brought to pH 5.7 with 6 N sodium deuteroxide and then filter sterilized. Optimal concentration of algal cell hydrolysate was determined by monitoring the growth of test cultures containing different concentrations of the hydrolysate.

4. Growth of Yeast in the Deuterated Medium

The yeast strain 20B-12 containing the overexpression plasmid pCGY219 was grown in deuterated media essentially as described by Mohan *et al.* (1962). The growth medium consisted of: yeast nitrogen base without amino acids (0.67%), adenine (0.01 mg/mL), uracil (0.01 mg/mL), thiamine pyridoxine, and myo-inositol (all to 0.5 mg/mL), 2,6 deuterated tyrosine (0.3 mg/mL), deuterated ethanol (2%) and algal cell hydrolysate (5%). Tyrosine was solublized in sodium deuteroxide and the pH of the media was set to 6.0 following the addition of the tyrosine. 750 mL of media was prepared. Growth was started by innoculating a 10 mL culture of protonated media. This culture was grown

overnight at 30° C. An aliquot of this culture was then added to a 50 mL aliquot of the deuterated medium, giving a final concentration of cells in the 50 mL aliquot of 4 x 10^6 cells/ mL. This culture was grown at 30° C with shaking in a water bath and the cell count was monitored daily. When the cell count reached 1 x 10^7 cells/ mL, the culture was diluted into the remaining 700 mL of deuterated medium. Again, cell count was monitored daily and cells were harvested in late log phase by centrifugation at 3000 x g for 15 minutes. Cells were then washed with sterile water and stored at -80° C until used. Growth of cells to a harvestable level required approximately a 1 month growth period.

5. Isolation of ^{1}H Tyr dPGK

The protocol used for the isolation of dPGK is essentially that of Kuntz and Kreitsch (1982). A brief summary of that procedure is as follows: The cells are subjected to three successive rounds of freezing at -80° C for at least one hour followed by thawing at room temperature for at least one hour. Following the freeze-thaw cycle, the cells are suspended in a lysis solution of equal volume containing 1 mM PMSF, 3 mM EDTA and 0.5 M ammonia, and water to volume. The final pH of this solution is approximately 10-11. The cells are incubated in the lysis solution for about eight hours with slow stirring. This is followed by incubation overnight without stirring. The cells are then brought to pH 7.5 with 5 M lactic acid. Cell debris is then separated by centrifugation at 30000 x g for one hour. The supernatant was collected and its volume determined. A 20 μL aliquot was removed for assay purposes. The pellet was saved for

possible further purification. The supernatant was again centrifuged for one hour at 20000 x g to remove any residual cell debris. The supernatant was collected and the volume determined. Another 20 μL aliquot was taken for assay purposes. Protamine sulfate was then added to the supernatant to a concentration of 0.1 g/ 100mL supernatant. This solution was then stirred at room temperature for 30 minutes and then centrifuged at 20000 x g for one hour. The supernatant from this centrifugation was then removed, the pellet saved, and the volume determined. Again, a 20 μL aliquot was taken for assay. DNAse and RNAse were then added to the supernatant at concentrations of 15 μg/mL and 20 μg/ mL, respectively. This solution was stirred for 15 minutes at room temperature. Ammonium sulfate was then added over a 15 minute period to a concentration of 30 g ammonium sulfate / 100 mL supernatant. This mixture was stirred 30 minutes at room temperature and then centrifuged at 20000 x g for 15 minutes. The supernatant was decanted and the pellet saved. A 20 μL aliquot was removed at this point in the prep. A second addition of ammonium sulfate was then made, at a concentration of 13 g ammonium sulfate / 100 g supernatant. The salt was added slowly over 15 minutes and allowed to stir for 30 minutes at room temperature. This solution was centrifuged at 20000 x g for 15 minutes. The supernatant from this centrifugation was decanted, an aliquot taken for assay and the pellet saved. The pH of the supernatant was then brought to 4.3 using 5 M lactic acid to precipitate the protein. This solution was then centrifuged at 20000 x g for 15 minutes. The supernatant was decanted and the pellet resuspended in 5 mM deuterated acetate buffer at pH 5.9. The suspension was then passed over a Sephacryl HR-200 column equilibrated with deuterated acetate buffer

at pH 5.9. Fractions containing protein were identified spectrophotometrically. Those fractions exhibiting an A_{280} / A_{260} ratio of 1 or greater and an A_{280} of 0.3 or greater were pooled. These fractions were lyophilized to dryness and resuspended in deuterated acetate buffer. Purity was determined by SDS-PAGE. Aliquots taken for enzyme activity assay were assayed immediately after being removed using the coupled enzyme assay described below.

6. Coupled Enzyme Assay To Determine PGK Activity

The reaction PGK catalyzes is most favorable in the glycolytic direction, generating ATP from 1,3-bisphosphoglycerate and MgADP. However, 1,3- bisphosphoglycerate is very unstable under the conditions of the assay, Therefore, most researchers assay PGK activity in the gluconeogenic direction. In this approach, 3-phosphoglycerate is incubated with MgATP, NADH, glyceraldehyde-3-phosphate dehydrogenase and the sample to be assayed (Bucher, 1942). The reaction that occurs is as follows:

$$\text{3-phosphoglycerate + MgATP} \xrightarrow{\text{PGK}} \text{1,3-bisphosphoglycerate + MgADP}$$

$$\text{1,3-bisphosphoglycerate + NADH} \xrightarrow{\text{GAPDH}} \text{glyceraldehyde-3-phosphate + } NAD^+$$

The activity of PGK is determined by monitoring the decrease in the absorbance of NADH at 340 nm in a spectrophotometer. 6220 $M^{-1}cm^{-1}$ was taken as the molar extinction coefficient, and the linear part of the trace is used for calculation of activity.

7. Nuclear Magnetic Resonance Studies

Protein samples used in NMR studies were prepared by repeated solublization and lyophilization in deuterium oxide. Sample volumes were from 400 to 500 μL. 2'-dATP was prepared as described above.

One-dimensional NMR Spectra

All NMR spectra in this section of the work were acquired on a wide-bore Bruker AMX spectrometer operating at 400 MHz. 32K data points were acquired, and a 90 degree observation pulse, typically 6.55 μsec, was used. Quadrature phase detection was used. Spectra were recorded at 300° K and processed with 1 Hz line broadening.

Substrate titrations were performed by sequentially adding 2'-dATP, $MgCl_2$, and glycerol-3-phosphate directly to the NMR tube. Substrates were added in three separate additions corresponding to substrate-to-enzyme ratios of 0.5:1, 1:1 and 1.5:1. $MgCl_2$ was added at the same time as the 2'-dATP. Glycerol-3-phosphate was added after the three additions of $MgCl_2$/2'-dATP. Following each addition, the sample was mixed by pipeting and a spectra taken. For each addition, 64 scans were taken.

Two-dimensional NMR Spectra

Nuclear Overhaüser Enhancement Spectroscopy (NOESY) experiments were performed on the ternary complex. The sample contained ^1H Tyr dPGK, $MgCl_2$, and glycerol-3-phosphate at a substrate-to enzyme ratio of 1.5:1. Spectra were acquired using the time proportional phase increment (TPPI) methodology (Marion & Wüthrich, 1983).

The mixing time for this experiment was 200 msec. This mixing time was varied by 5%. The spectral width in these experiments was 4032 Hz. 2K data points were recorded for each of 180 slices. Each slice consisted of 216 scans. The data were zero filled to 1K, and a $\sin^2$ window function was applied prior to Fourier transformation.

C. Results

1. Characterization of 2,6 Deuterated Tyrosine

The results of the selective deuteration of tyrosine are shown in Figure II-4. Figure II-4 shows the proton spectrum of both isotopically normal tyrosine and the proton spectrum of 2,6 deuterated tyrosine. It is clear from the figure that deuteration was successful and the resulting singlet, located at a position between the two peaks in the isotopically normal amino acid, indicates that the splitting was indeed removed.

2. *Saccharomyces cerevisiae* Growth on the Deuterated Medium

The growth of *S. cerevisiae* in the deuterated media was much slower than that observed with yeast growing in isotopically normal growth media. The cells appeared morphologically normal. They did not reach the levels of growth obtainable from isotopically normal media. In this growth cell density reached only 1 $x10^7$, a full order of magnitude lower than that routinely achieved by growth under normal conditions. This effect was also observed by Shibata *et al.* Protonated tyrosine was added to a concentration of 0.035%.

3. Isolation of ^{1}H Tyr dPGK

The results of activity assays taken during the purification of the enzyme (data not shown) indicate that there was minimal loss of protein during the prep. A final yield of 8.2 mg of protein was achieved from the 1L culture. SDS-PAGE showed that the enzyme was free of contaminants (data not shown). This yield is significantly lower than that achievable from normal protein preps. The phenomenon of lower protein production in the deuterated system has been previously observed (Pappu, 1996).

4. NMR Characterization of the ^{1}H Tyr dPGK Enzyme

Figure II-5 shows the one-dimensional ^{1}H NMR spectrum of the 3,5^{1}HTyr dPGK enzyme (referred to hereafter as the "tyrosine-labeled enzyme"). A number of features of this spectrum are of import. The most striking feature of the spectrum is the clear background as a result of deuteration. The entire envelope of aromatic peaks has been reduced to a single group corresponding to the tyrosine ring protons. Note that the aromatic region of the spectrum is completely free of any contaminating aromatic or amine resonances. Further expansion and analysis of the aromatic region (not shown) reveal that there are no other residual peaks in this region of the spectrum and that the deuteration is complete. This analysis also showed that the tyrosine ring peak did not show any splitting due to the presence of 2,6 protons.

The αH peaks of the tyrosine residues are seen as a sharp multiplet centered at 4.2 ppm. These peaks are not identifiable in the fully protonated enzyme. A comparison with the spectrum of the fully protonated enzyme with the tyrosine-labeled enzyme emphasizes this aspect. Again, the simplification of this region of the spectrum due to deuteration is obvious. Seven peaks are expected in this region, although overlap of the resonances does not permit a reliable estimation of the peak number.

The βH peaks of tyrosine are also seen as a group centered around 3.8 ppm. These peaks are somewhat more separated than the α peaks. There are seven peaks expected in this region. Approximately seven peaks are observed in this region.

There are two large peaks centered around 1.5 ppm in the spectrum. These peaks do not correspond to tyrosine resonances. A comparison of the chemical shifts of these peaks with the known chemical shifts for the twenty amino acids suggests that the peaks are due to the presence of protonated alanine in the enzyme. The peak at 1.5 ppm corresponds to the βH protons of alanine. This scrambling of the label is a result of the metabolization of the tyrosine label and subsequent reincorporation of the atoms of tyrosine into alanine. The biosynthesized alanine is then reincorporated into the growing enzyme. The key enzyme in the pathway that converts tyrosine into alanine is a transaminase. This problem of reincorporation of the tyrosine label as alanine was a recurrent problem in the deuterated preparations of the various labeled enzymes discussed in this dissertation. This problem was identified in the preparation of the tyrosine enzyme and was addressed in future preparations.

5. Substrate Titrations of the Tyrosine-Labeled Enzyme

Following initial characterization of the enzyme, substrate titrations were performed. The enzyme was first titrated with 2'-dATP and $MgCl_2$, together, to give the same substrate-to-enzyme ratio. 2'-dATP was chosen for this work because the 2'H and 2"H resonances are shifted upfield and are not subject to being obscured by the peak due to residual water. It is known that 2'-dATP is a substrate for PGK.

Substrate-to-enzyme ratios of 0.5:1, 1:1 and 1.5:1 were used, and spectra were taken directly after addition of the substrate. At the lowest substrate-to-enzyme ratio of MgATP:^{1}H Tyr dPGK (Figure II-6) there were a number of changes in the proton spectrum.

The first feature to note in the spectrum of the enzyme in the presence of Mg 2'-dATP (Figure II-6) are the peaks corresponding to the H2' and H2" protons. These peaks appear in the spectrum at 2.6 ppm and 2.8 ppm, respectively. These peaks are not visible in the fully protonated enzyme. In the tyrosine-labeled enzyme they are clearly visible, well-resolved and even the splittings of the peaks are visible. This fact further emphasizes the advantages of using the deuterated enzyme over the protonated enzyme in NMR studies. It is likely that the use of a partially deuterated or a fractionally deuterated labeling strategy would not be able to produce the same quality of spectrum as that of Figure II-6. Use of these methods would increase the number of magnetically sensitive nuclei in the protein and as a result, the number of relaxation pathways would increase. This would lead to a broadened line shape and a decrease in peak intensity.

There are three changes in the spectrum of the tyrosine-labeled enzyme that occur upon the addition of Mg 2'-dATP. These changes are presented in the lower trace of Figure II-6 and the affected peaks are labeled as a, b and c. The lower trace represents the enzyme in the absence of substrates. This figure shows an expansion of the aliphatic region of the spectrum. Peak a is seen to disappear, due to line broadening. Similarly, peak b is also seen to disappear. Peak c is seen to shift its position from about 3.3 ppm to about 3.2 ppm. Peaks a and b are likely α protons of tyrosine and peak c is likely a β proton of tyrosine. These changes in the spectrum are due to a change in the local environment of these protons upon addition of Mg 2'-dATP. Further addition of Mg 2'-dATP did not alter the spectrum further.

The substrate titration was continued by the addition of the competitive inhibitor, glycerol-3-phosphate (G3P). Three additions were made to give substrate-to-enzyme ratios of 0.5:1, 1:1 and 1.5:1. This addition represents the formation of the ternary complex. Figure II-7 shows the results of these additions. None of the three additions produced any further change in the spectrum.

6. NOESY Spectrum of the Tyrosine-Labeled Enzyme in the Presence of Substrates

Following the addition of all substrates, a NOESY experiment was performed on the complex (Figure II-8). This spectrum was recorded with a mixing time of 200 msec. The first thing that is apparent in this spectrum is the lack of intraprotein crosspeaks. This clear background allows identification of peaks that are not observable in NOESY spectra of the fully protonated enzyme (Shibata *et al.*, 1995).

The NOESY spectrum shows three crosspeaks of significant importance. These peaks are labeled in Figure II-8. One peak identified in the spectrum corresponds to a crosspeak between the H2' and H2" protons of 2'-dATP. This crosspeak is difficult to observe in the protonated enzyme due to spectral overlap. The ability to clearly identifiy this peak in Figure II-8 is indicative of the effectiveness of the deuteration process.

The second labeled peak corresponds to a NOE interaction between the 2' proton of 2'-dATP and the β proton of one of the tyrosines. This enzyme-to-substrate interaction implies that at least one of the tyrosine residues is within 4 Å of the ribose ring.

The third labeled peak shows the expected NOE interaction between the α and β protons of the tyrosine residues. No other intraresidue crosspeaks are observed, as expected. A labeling with fully protonated tyrosine would be expected to produce an interaction between the 3,5 and 2,6 ring protons, but the use of the 2,6 deuterated tyrosine in this preparation eliminates the potential for that interaction.

D. Discussion

There are two important results of this work; firstly, the work demonstrates the feasibility of deuterating a large protein and of incorporation of protonated labels at known positions in that protein. This labeling strategy allows detailed studies not possible using the isotopically normal enzyme. Secondly, substrate titrations reveal that at least one of the proton-labeled tyrosine residues is in close proximity to the 2' proton of

the ribose ring of 2'-dATP in the binary complex of Mg 2'-dATP and the tyrosine-labeled enzyme.

As discussed above, the study of proteins larger than 25 kDa using traditional NMR methods is not possible due to overlap of peaks and poor resolution due to the short relaxation times that accompany the large number of protons present. The problem of overlap is adequately solved using the process of deuteration of the enzyme. The only overlap that is seen in the proton spectrum of the tyrosine labeled enzyme is due to the overlap of the tyrosine residues themselves. A somewhat better dispersion of the ring and β protons was hoped for, but overlap was not unexpected. In order for the peaks to be separated in the spectrum, a shift in the local environment of the protons is required. The side-chains of residues are most susceptible to local environment and it is reasonable to expect some dispersion in their positions. The β protons and especially the α protons are much less likely to show separation, since their local environment is much more probable to be similar throughout the protein. It should be quickly noted that although this overlap was present, it in no way deterred the main goals of this set of experiments, which was to show the possibility for specific labeling and deuteration of a large protein for use in detailed structural NMR studies.

The reduction in protons as a result of the deuteration of the protein also leads to the lengthening of relaxation rates of both the proton labels of the protein and the substrates. This effect manifests itself in the proton NMR spectrum as a decrease in the line width of observed peaks. The individual peaks of the tyrosine residues and substrate protons are clearly visible and, as mentioned above, even the splittings of the substrate

peaks are visible. This resolution is not possible in the protonated enzyme due to the multitude of available relaxation pathways. That the substrate peaks are visible at all is further evidence for the utility of the deuterated system.

The substrate titration data suggest that at least one tyrosine is affected by the addition of Mg 2'-dATP. The disappearance of the two peaks implies that their local environment is changing. The same is true of the third peak, which shifts position. It is interesting to note that addition of G3P did not alter the spectrum, since it has been reported that the binding of PGA causes significant structural changes in the enzyme (Harlos *et al.*, 1992). It possible that the binding of G3P did, in fact, induce conformational changes in the enzyme, however the labeled tyrosine residues were unaffected by this change. If a different residue of the protein was isotopically labeled, the conformational change induced by binding G3P may be reflected in a change in intensity or position of these probe residues. Protonated histidine residues were incorporated into deuterated PGK in a similar study (Pappu & Serpersu, 1994). It is known that several histidine residues are located near the positively-charged PGA binding site. In that study, several peaks were observed to change position when G3P was added, consistent with the proximity of these histidines to the PGA binding site. Therefore, it is not surprising that the tyrosine labels did not shift position, as they are not located near the PGA binding site.

From the substrate data alone, it is not possible to determine what types of conformational changes are occurring in the protein to bring about these alterations in the spectrum. In the current work, all seven tyrosine residues of PGK were proton labeled,

and it is not possible to tell which tyrosine residue corresponds to which peak in the spectrum. In order to make absolute statements about which tyrosine residues are affected in which manner, it is necessary to design a method of labeling just one or two of these tyrosines in such a way as to allow their identification in the NMR spectrum. This problem is addressed in the next chapter of this dissertation.

The NOESY experiment provided further structural information about the ternary complex. One NOE interaction was observed between the 2' proton of 2'-dATP and a β proton of one of the tyrosine residues. Again, not having specifically labeled the tyrosine residues, it is not possible to definitively assign the β proton to any one of the seven residues. These data also suggested that the labeling of tyrosine residues for use as structural probes was sound, and that further labeling studies involving this residue should be pursued.

Chapter III

Site-Specific Labeling of Deuterated PGK

A. Introduction

The problems involved with the study of large proteins using conventional NMR studies have been addressed in the previous chapter, which described work in which PGK was biosynthetically prepared in a fully deuterated condition, with the exception of the seven tyrosine residues of PGK. These tyrosine residues were protonated at all positions except the 2,6 ring positions. This study provided general information about the involvement of the tyrosine residues in the Mg 2'-dATP binding event. However, the information was not "residue-specific". That is, the labeling strategy did not permit the assignment of any given one of the seven tyrosines, and only allowed general conclusions to be drawn about the nature of the interaction between the β proton of a tyrosine and the ribose ring of 2'-dATP. What is required to complete this study is the development of a method that will allow identification of the specific tyrosine(s) that is involved in this interaction. Analysis of the previous NMR data in conjunction with data from crystallographic studies may shed some light on the identity of the residue(s) involved in this interaction, but such an analysis would still be speculative. What is required is a method by which one or two residues of a large protein such as PGK can be isotopically labeled and subsequently identified in the NMR spectrum. As detailed below, one method for site-specific identification of a given residue in a protein involves isotopically

labeling two sequential residues. These isotopically labeled residues can then be correlated using two and three dimensional NMR experiments. The development of a protocol to allow such a labeling is the focus of this chapter.

The sequential resonance assignment of the proton NMR spectrum of a protein is an excellent way of obtaining structural information about the protein. If each peak in the proton spectrum is known, analysis of one and two dimensional experiments can yield detailed information about the state of a protein under different sets of conditions. The ability to make sequential assignments has long been a goal of researchers in the area of protein NMR, but the size of the protein under study has been a limiting factor. As the size of the protein increases, the number of protons in the spectrum increases accordingly. This leads to complicated, unanalyzable spectra for proteins above 25 kDa.

The first complete protein NMR sequential assignment was reported in 1982 (Wagner & Wüthrich, 1982). This study focused on the resonance assignments of basic pancreatic trypsin inhibitor (BPTI), a small (58 residue) protein of molecular weight 6500 Da. This study used the two dimensional NMR techniques COSY and NOESY, in conjunction with one dimensional spectra to solve the resonance assignments of this protein. The COSY (COrrelated SpectroscopY) experiment relies on through-bond coupling. These three-bond interactions are represented as crosspeaks in the spectrum. NOESY (Nuclear Overhaüser Enhancement SpectroscopY) experiments rely on through-space interactions. Magnetization is generated on one nucleus and transferred through space, within a 4 Å radius, to neighboring nuclei. In Wagner & Wüthrich's study, COSY experiments were used to identify spin systems of the individual residues

and NOESY experiments were used to link the residues in sequence. This method works well for peptides and small proteins, however for larger proteins, the extensive overlap of peaks requires the use of other methodology.

Another approach to NMR studies of large proteins is to use isotope labeling. In this strategy, the protein is typically uniformly labeled with ^{15}N, ^{13}C or partially labeled with ^{2}H (LeMaster & Richards, 1988). These labeling strategies can help alleviate overcrowding of peaks in the two dimensional spectrum. However, even with these isotope editing approaches, sequential assignments of large proteins are difficult and rely on NOE experiments. With such a heavy dependency on the NOESY experiment, spin diffusion, backbone conformation and other experimental concerns can make analysis of such spectra almost impossible. Although the advent of three dimensional experiments (^{15}N, ^{13}C and ^{1}H) has greatly simplified spectra, researchers have yet to be successful in completely assigning the resonances of a large protein. However, partial assignments have been successfully made (Ikura *et al*, 1990) using this method.

Triple resonance studies have been successful in assigning the backbone resonances of several proteins (Ikura *et al*, 1990; Yamazaki *et al*, 1997). These experiments do not rely on NOESY experiments and rather focus on pulse programs to take advantage of unique labelings. Proteins are uniformly labeled with ^{15}N and ^{13}C, at the α position of each residue. Residues are then linked by experiments that correlate the αH to the ^{15}N and ^{13}C labels. Again, this method has been applied to proteins smaller than PGK, although the technology has been used to study protein complexes of up to 64

kDa (Shan *et al*, 1996). A further problem with these triple resonance studies of large proteins is that the number of protons leads to faster transverse relaxation and therefore a decrease in the resolution of peaks in the spectrum. The efficiency of magnetization transfer is also impaired by the presence of this large volume of protons.

Some researchers have attempted to avoid these problems using the various forms of deuteration discussed above (Chapter II, Introduction). These experiments have relied on nonselective deuteration of the protein under study. Nonselective deuteration techniques have proven to be useful for making backbone structural assignments for smaller proteins, but to date, no technique has been perfected for performing resonance assignments for large proteins that have protons attached to an $^{13}C_{\alpha}$ label.

The methodology developed in this laboratory, as part of this dissertation, is directed at demonstrating the feasibility of labeling strategies for large proteins that lead to site specific resonance assignments. The protein used in this study takes advantage of the location of two of the tyrosine residues in the primary sequence of PGK. The seven tyrosine residues are found at the following positions in the primary sequence of PGK:

Tyr 48-Val 49; Tyr 56-Val 57; Tyr 74-Ser 75; Tyr 122-His 123;
Tyr 158-Ile 159; Tyr 193-Phe 194; Tyr 380-Gly 381;

Note that there are two instances in the sequence where a tyrosine is adjacent in sequence to a valine residue. The labeling strategy used in this work takes advantage of these two occurrences; the protein is grown in the presence of both $^{13}C_{\alpha}$ tyrosine and ^{15}N valine. The protein is also grown on a deuterated media to yield a PGK protein that

is perdeuterated, with the exception of the tyrosine and valine residues, and has an $^{13}C_{\alpha}$ label at each tyrosine and an ^{15}N label at each of the valine residues. There are 37 valines in PGK and the ^{15}N label is expected to be incorporated at each position, however this incorporation has no adverse effect on the NMR experiments performed; the presence of the ^{15}N label at positions in the protein other than those of primary interest will not interfere with interpretation of the HNCA experiment, since this experiment is based on the correlation of the ^{15}N label and a sequential $^{13}C_{\alpha}$ label. The labeling places a protonated $^{13}C_{\alpha}$ tyrosine next to a protonated ^{15}N valine at positions 48 and 56. These two residues can be correlated using three dimensional HNCA experiments and the connectivity traced from each label to its side chain protons using various two dimensional NMR techniques.

A similar PGK protein was prepared that used the same approach to allow identification of tyrosines 122, 158 and 193. In this preparation, the labels were protonated $^{13}C_{\alpha}$ tyrosine and both protonated ^{15}N labeled phenylalanine and protonated ^{15}N labeled isoleucine. Isoleucine appears in the primary sequence of PGK at positions 123 and 159. Phenylalanine appears at position 194. The ^{15}N isotope labels can be correlated with the $^{13}C_{\alpha}$ tyrosine labels using the methods discussed above.

This unique labeling strategy represents the first attempt to identify, site specifically, two resonances of a large protein by perdeuteration and specific labeling. This method does not rely on NOESY experiments and is singular in that it utilizes fully

protonated labeled amino acids. Successful identification of these resonances will allow further experiments designed to probe conformational changes of the enzyme in the presence of substrates. Since the labels are themselves protonated, NOESY experiments can be performed on the enzyme-substrate complex to learn more about the molecular motions undergone by this protein upon binding substrates.

B. Experimental Procedures

1. Materials

Materials used were obtained from the same sources and prepared in the same manner as reported earlier (Experimental Procedures, Chapter II). $^{13}C_{\alpha}$ tyrosine, ^{15}N valine, ^{15}N isoleucine and ^{15}N phenylalanine were purchased from Isotec, Inc.

2. Yeast Growth in Deuterated Medium

S. cerevisiae strain 20-B12 containing multicopy plasmid pCGY219 was grown as previously described (Experimental Procedures, Chapter II).

In the ^{15}N valine preparation, 0.2% deuterated glycerol was included in the medium as the carbon source, instead of algal cell hydrolysate. The reason for this change will be discussed below. The medium was the same as that used in isotopically normal growths (Scopes, 1971).

In the ^{15}N phenylalanine/ ^{15}N isoleucine preparation, 0.2% deuterated glucose was used as the carbon source. The medium used in this growth consisted of $MgSO_4 \cdot 7H_2O$ (0.07%), KH_2PO_4 (0.1%), $CaCl_2$ (0.04%), NaCl (0.05%), $(NH_4)_2SO_4$ (0.12%), Citric Acid (1.05%), 10 M KOH in deuterium oxide (1.6%) and "Trace Metals" (1.0%) and deuterium oxide to volume. The pH was set to 6.1, filter sterilized and 5 mL of vitamin solution in deuterium oxide was added.

Third, at least two labels were included in the growth medium, $^{13}C_\alpha$ tyrosine (100 μg/mL) and either ^{15}N valine (100 μg/mL),^{15}N isoleucine (100 μg/mL) or ^{15}N phenylalanine (100 μg/mL). No selectively deuterated tyrosine was used.

The protocols for cell growth discussed above were followed and the cells were harvested, processed and stored in the same way as previously described.

3. Isolation of $^{13}C_\alpha$ Tyrosine ^{15}N Isoleucine ^{15}N Phenylalanine dPGK and $^{13}C_\alpha$ Tyrosine ^{15}N Valine dPGK

Both enzymes were prepared using the method of Kuntz and Kreitsch (1982), as described above (Experimental Procedures, Chapter II) with the following exception: In the $^{13}C_\alpha$ tyrosine ^{15}N valine dPGK preparation, the pH 4.3 precipitate was not applied to the Sephacryl column in the interest of preventing loss or dilution of the protein. This precipitate was resolublized in 5 mM deuterated acetate buffer, pH 5.9 and an aliquot was analyzed by SDS-PAGE. It was determined that the enzyme was of sufficient purity

for NMR experiments. The A_{280}/A_{260} ratio was taken to determine the extent of nucleotide present. The sample was then treated with DNAse and the A_{280}/A_{260} taken again. It was determined that the sample was adequately free of nucleotide contamination and was stored at 4° C in 5 mM deuterated acetate buffer. Activity was assayed over the course of the preparation using the coupled enzyme assay described above (Experimental Procedures, Chapter II).

4. Magnetic Resonance Studies

Samples were prepared for NMR experiments by repeated lyophilization and resolublization in deuterium oxide. Final volume for the experiments was 500 mL. Samples were used immediately after the final lyophilization/resolublization cycle.

4.1 *One and Two Dimensional NMR Experiments*

All one dimensional NMR experiments were performed on a wide bore Bruker AMX spectrometer operating at 400 MHz. All spectra were acquired at 300° K and 32K data points were collected. Chemical shifts were referenced to an external DSS sample. Spectra were processed with 1 Hz line broadening. The X-filtered experiment was performed at the NMR Facility at Madison, Wisconsin (NMRFAM) on a Bruker DMX spectrometer operating at 500 MHz. The concentration of protein in the sample was about 0.5 mM.

4.2 *Two and Three Dimensional NMR Experiments*

Two dimensional NOESY, inverse-detected $^{13}C_\alpha$-H_α correlation (INV4TP) (Bax *et al*, 1983) and inverse-detected, MLEV-17-based $^{13}C_\alpha$-H_α-H_β correlation (INV4MLNDTP) (Lerner *et al*, 1986) experiments were performed on the Bruker AMX spectrometer described above.

Triple resonance ^{1}H-^{15}N-$^{13}C_\alpha$ correlation (HNCA) (Kay *et al*, 1990) and heteronuclear single quantum coherence (HSQC) (Bodenhaüsen & Ruben, 1980) experiments were performed at the NMR Facility at Madison, Wisconsin (NMRFAM) on a Bruker DMX spectrometer operating at 500 MHz. The HNCA pulse program was modified somewhat to allow expression of the data in two dimensions. The program does not collect data in the third ^{15}N dimension, but rather treats this dimension as a filter and shows only the carbon to proton connectivities. This editing makes the HNCA experiment essentially a filtered two dimensional experiment.

C. Results

1. <u>$^{13}C_\alpha$ Tyrosine ^{15}N Isoleucine ^{15}N Phenylalanine dPGK Growth in The Deuterated Medium</u>

Yeast were force-fed isotope-labeled amino acids in both of the labeled dPGK preparations discussed. It has been demonstrated in this work (Chapter II) and previously (Crespi *et al*, 1972) that force-feeding is an effective method for preparing isotope

labeled proteins. In the first preparation, $^{13}C_{\alpha}$ tyrosine, ^{15}N isoleucine and ^{15}N phenylalanine were included in the growth media. Labels were included at a concentration of 100 mg/mL, which was determined in the ^{1}H tyrosine dPGK preparation to be the optimal concentration, as determined using a series of test cultures.

This culture was also grown using deuterated glucose as the carbon source. This procedure replaced the use of deuterated algal cell hydrolysate, which was used in all previous preparations. The decision to switch to the use of deuterated glucose was made when the quality of available deuterated algal cells dropped drastically. Hydrolysate from these deuterated algal cells did not produce acceptable levels of growth or satisfactory morphology of the yeast cells in test cultures. Several batches of deuterated algal cells were used, each with the same unacceptable results.

The method of preparation of the algal cell hydrolysate was also modified. Silver carbonate, rather than sodium deuteroxide, was used to bring the pH of the cells to pH 5.7. This modification did not alter the pattern of low cell counts and small yeast cell size, as viewed through a phase contrast microscope. Other aspects of the hydrolysis were also modified. These modifications include varying the concentration of deuterium chloride used in the hydrolysis, varying the time of refluxing, and varying the temperature at which the cells were hydrolyzed. The effect of each of these modifications was assessed by monitoring test cultures. No improvement in cell growth was observed for any of these changes to the standard procedure. Fresh slants were also prepared and used to start the cultures, but this did not increase the cell counts.

As a result of this study, it was decided that another source of carbon should be investigated. Deuterated glucose was selected as the carbon source because it had been used successfully in earlier studies (Scopes, 1971). The optimal concentration of deuterated glucose in the growth media was determined by monitoring the growth of small test cultures. It was determined that 0.2% was the optimal concentration for growth, although the cells counts were lower than those achieved using deuterated algal cell hydrolysate. A number of growth media were investigated by monitoring yeast cell growth and the medium described above was chosen. Again, this medium has been used successfully in other experiments.

Following this extremely time-intensive survey of conditions, the $^{13}C_{\alpha}$ Tyr ^{15}N Ile ^{15}N Phe dPGK growth was begun. The yeast cells grew much more slowly than they did using the deuterated algal cell hydrolysate, and the growth of the 750 mL culture took over three weeks to grow to a level suitable for harvesting. They also did not achieve the same cell count observed in either normal or earlier deuterated preparations. Isolation of the labeled enzyme from this growth yielded only 6 mg of protein. This amount is somewhat lower than that obtained from other deuterated preparations, but similar to the yield from other doubly labeled deuterated growths in our laboratory (Pappu, 1996).

2. Characterization of The $^{13}C_{\alpha}$ Tyr ^{15}N Ile ^{15}N Phe dPGK Enzyme

The one dimensional proton NMR spectrum of the $^{13}C_{\alpha}$ Tyr ^{15}N Ile ^{15}N dPGK enzyme is shown in Figure III-1. From this spectrum it can be seen that the labels did

incorporate successfully. The ring protons of the tyrosine are visible, as are the ring protons of the phenylalanine. The αH and βH peaks of the labels are also observed.

The spectrum initially appears somewhat crowded with protons. However, it must be noted that there are 23 isoleucines and 19 phenylalanines in PGK, as well as the seven tyrosines. This is a total of 49 protonated residues, giving a total of approximately 400 protons. Although this number may seem high, and the spectrum somewhat more complicated, the important point is that this represents a reduction in the number of peaks found in the spectrum of isotopically normal PGK by a full order of magnitude.

Another important aspect of this spectrum is the completeness of deuteration. On the downfield side of water, the trace returns to baseline and remains there until the peaks of the aromatic residues appear. A comparison of this spectrum with the isotopically normal enzyme shows the simplification achieved by deuteration of the enzyme. The effect of the deuteration is also manifested in the narrow line widths and sharp peaks of the spectrum. As discussed earlier, this is a result of the lengthening of the relaxation times by the removal of protons and their replacement with deuterons, which have a much smaller magnetic moment.

The ^{13}C spectrum of this enzyme was then taken. It is shown in Figure III-2, as a full spectrum. The spectrum is expected to show seven peaks, corresponding to the seven $^{13}C_{\alpha}$ tyrosine labels. Peaks are seen around 170 ppm, 35 ppm and 20 ppm, which indicates some degree of label scrambling. However, peaks are also seen around 50 ppm, which is where the α carbons are expected to appear. There are seven expected peaks in this region of the spectrum, but it is difficult to identify the expected peaks. This is due

to the low concentration of the sample, which was about 0.25 mM in the NMR tube. Also, this experiment utilizes direct detection of ^{13}C, which is a less sensitive nucleus, compared to a proton: if the sensitivity of a given number of protons is assigned a value of 1, the relative sensitivity of an equal number of ^{13}C spins is 1.59×10^{-2} (i.e $(\gamma_C/\gamma_H)^3$) (Sanders & Hunter, 1994). This decreased sensitivity is reflected in the spectrum of the enzyme.

Since the concentration of the sample was low, the remaining experiments were not feasible to perform on the AMX 400 spectrometer at UTK. Acquisition of the required ^{15}N spectra would not be practical due to both sensitivity and time constraints. Moreover, the three dimensional experiments require an instrument with three channels and this instrumentation was not available at UTK. Therefore, the sample was taken to the National NMR Facility at Madison (NMRFAM).

At NMRFAM, an X-filtered experiment was performed. This spectrum is shown in Figure III-3. In this experiment, protons are detected, however, the pulse program imposes the requirement that detected protons must be attached to an ^{15}N nucleus. The proton's magnetization is thereby "filtered" through the ^{15}N nucleus. Figure III-3 shows that the only signals that appear in the spectrum appear in the amine region of the spectrum. This peak is not well resolved, but it represents all 42 ^{15}N labeled nuclei, and therefore excellent resolution is not expected. The small non-phaseable peak around 4.8 ppm is likely a result of the presence of water and is not representative of labeling; it is

an instrumental artifact. Figure III-3 demonstrates that there is minimal, if any, scrambling of the ^{15}N label.

The next experiment performed at NMRFAM was an Heteronuclear Single Quantum Coherence (HSQC) experiment (Bodenhaüsen and Rubin, 1980). This experiment is a two dimensional experiment that correlates protons to an X nucleus, in this case, ^{15}N. The experiment uses an INEPT transfer (Insensitive Nuclei Enhanced by Polarization Transfer) to transfer magnetization from protons to ^{15}N. A 180 degree pulse on the protons is then applied, and a reverse INEPT transfer is used to transfer magnetization back from the ^{15}N nuclei to the proton. This experiment is essentially the two-dimensional version of the X-filtered experiment described in Figure III-3. Figure III-4 shows the two dimensional HSQC spectrum.

The most important piece of information to be learned from analysis of this spectrum is the following: as mentioned above, this spectrum represents connectivities between ^{15}N nuclei and their attached protons. There are a total of 42 such instances in the labeled protein. In the isotopically-normal enzyme, an HSQC experiment such as this would yield one crosspeak for each of the residues of the protein. This would give 416 crosspeaks in the same spectral region, a totally unanalyzable volume of signals. Figure III-4 represents a reduction in the number of peaks of a full order of magnitude and this region is still crowded and difficult to analyze. This spectrum emphasizes the need for further three dimensional experiments. Such experiments will serve to shift all but two of these peaks into the third dimension and allow conclusive identification of the

phenylalanine and isoleucine-labeled tyrosine residues. One of the goals of this trip to NMRFAM was to perform an HNCA experiment on this labeled protein, but due to instrumental difficulties, this was not possible at this time.

3. Growth of ^{15}N Valine $^{13}C_{\alpha}$ Tyrosine dPGK on Deuterated Glycerol

The low yield of protein and slow growth of the yeast cells in the deuterated glucose medium was a concern for later preparations. NMR experiments require a significant amount of protein, and it was not practical to simply repeat the growths over and over to obtain an acceptable amount of cells. It was decided then, that two different approaches to improving the yield of cells and amount of protein from each growth would be investigated. The first approach involved the expression of the protein in an *Escherichia coli* overexpression system. This approach will be discussed in Chapter IV. The second appraoch was to find an alternative carbon source that would allow the cells to grow faster and to greater numbers. To this end, deuterated glycerol was investigated as a possible alternative.

The medium used for the growth on deuterated glycerol was the same as is used in the regular PGK preparations (Scopes, 1971), however deuterated glycerol was used in place of glucose. A series of test cultures was prepared, in which the pH as well as the concentration of glycerol was varied. The optimal growth conditions were found to be pH 6.1 and 0.2% glycerol. These test cultures used isotopically normal glycerol, although deuterated glycerol was used in the labeled growths. The cells in the test

cultures appeared to be morphologically similar to those grown on glucose and grew to acceptable numbers (~1x 10^7).

The labeling strategy was also somewhat altered at this point. Rather than repeat the ^{15}N Ile ^{15}N Phe labeling, a new labeling strategy was devised. It was decided that instead of including two different ^{15}N labels, only one would be used: ^{15}N valine. The overall idea was the same: there are 38 valines in the sequence of PGK, but there are two instances where there is a valine located next to a tyrosine at the n+1 position. These instances are: Tyr 56-Val 57 and Tyr 48-Val 49. By including the α^{13}C tyrosine and ^{15}N valine labels, both of these tyrosine residues can be unequivocally identified using a similar series of experiments as those performed on the Ile-Phe-Tyr dPGK enzyme. It was expected that the two residues could then be distinguished from each other based on their chemical shifts, and consideration of their local environments.

Following optimization of growth conditions, the growth including ^{15}N valine was undertaken. The cells grew very slowly and again, over a 3 week growth period was required. Cell growth was monitored daily. Cells appeared to be healthy and morphologically normal. The cells were harvested when they reached a density of about 1 x 10^7 cells/mL. The medium was recovered, filter sterilized, additional deuterated glycerol added, and used in other growths. The cells were pooled and stored at -80° C. It is interesting to note that subsequent growths proceeded at faster growth rates than the initial growth on this medium. This reduction in growth time may be attributable to the

presence of growth factors or other compounds released by the yeast in prior growths. Approximately 15 g (wet weight) of cells were recovered from three different growths.

Isolation of enzyme from these cells gave only about 8 mg of labeled enzyme, which was lower than expected. The preparation of the enzyme was monitored by activity assay during the course of the isolation and there was no decrease in the level of activity. This suggests that there was no loss of enzyme during the isolation procedure, rather the expression level of the protein in this system is low. This research represents the first use of deuterated glycerol as a carbon source for yeast PGK deuteration and labeling, and it may be that the expression level is compromised by the growth conditions. Again, low protein yield from deuterated preparations is not uncommon (Pappu, 1996), and ultimately, this low yield of protein did not prevent the completion of any of the NMR experiments proposed.

4. Characterization of The ^{15}N Val $^{13}C_{\alpha}$ Tyr dPGK Enzyme

The one dimensional spectrum of the valine enzyme is shown in Figure III-5. The expected features of this spectrum are clearly visible: the α protons appear around 4.0 ppm and represent a total of 45 protons. They are grouped together and although there is some degree of overlap, many appear as sharp peaks. This feature is not achievable in the protonated enzyme. Some degree of overlap is expected, since the α proton region spans only about one ppm and contains 45 signals. However, it is important to note that the envelope in this region drops to the baseline on either side of this group of peaks.

The β protons of tyrosine appear around 3 ppm. They are closely grouped together and comprise a single intense signal. The β protons of valine appear further upfield around 2 ppm. There are a total of 45 expected signals in this region, so a degree of overlap is expected and observed. However, many of the peaks are sharp and well-resolved.

The methyl peaks of valine are also present in this spectrum and seen around 1 ppm. It is likely that there is some overlap between the β and methyl signals of the enzyme, both above and below 1 ppm. This overlap arises due to similar local environments. The variety of local environments present in this enzyme is exemplified by the appearance of a peak at around 0 ppm. This peak represents valine methyl groups that have shifted upfield due to an increase in the hydrophobicity of the local environment of these residues.

The tyrosine ring protons are observable around 6.9 to 7.6 ppm. These peaks represent 14 protons. Several sharp peaks are seen around 8 ppm. These peaks are due to the presence of protonated amines. Also note the complete absence of peaks in the 6.9-5 ppm region, a result of efficent deuteration of the protein. The intense peak around 5 ppm corresponds to the residual water resonance.

The ^{15}N Val $^{13}C_{\alpha}$ Tyr dPGK enzyme was then studied using a Total Correlation Spectroscopy (TOCSY) experiment (Bax & Davis, 1985). This experiment effectively extends the capabilities of a Correlated Spectroscopy (COSY) experiment (Jeener, 1971). The homonuclear COSY experiment correlates protons based on a geminal or vicinal spin coupling, while the TOCSY experiment uses an MLEV-17 spin-lock sequence to

relay magnetization from one nucleus to the next, in a scalar-coupled spin system (Braun *et al*, 1996). The TOCSY experiment, then , is used when visualization of all possible scalar connectivies of a system is desired. The variable in this experiment is the length of the spin-lock: shorter spin-locks give essentially the same connectivity information that a COSY experiment does; longer spin-locks provide the total scalar connectivity of a sample.

Figure III-6 shows the TOCSY spectrum of the valine-tyrosine labeled dPGK. This figure represents only one region of the spectrum. In this region, only the αH-to-βH connectivies of tyrosine are observed. Seven peaks are expected in this region, and approximately seven peaks are observed. The αH-to-βH connectivities of the valine residues are also observed in this spectrum, but are not observed in the region shown in Figure III-6. The βH resonances of valine are found further upfield around 2 ppm, and therefore are not seen in the same area as the peaks of the tyrosine resonances. The TOCSY experiment relies on sequential, scalar connectivities. As such, there is no connectivity observed between the ring and β protons of tyrosine. In the structure of this residue, the β protons are separated from the ring protons by an unprotonated carbon, which cannot accept the magnetization transferred during the spin-lock, and correlation of the system stops at the β proton.

The next step in the characterization of this enzyme was the examination of the carbon spectrum of the enzyme. This spectrum showed a very low signal to noise ratio, and virtually no useful information about the enzyme can be extracted from this spectrum. This was likely due not only to the low concentration of the sample, but to

substantial exhange of deuterons for alpha-protons. It is possible to decouple the deuterium from the ^{13}C, however this requires a highly specialized piece of hardware not available on site. This spectrum did give an indication of the degree of exchange that occurred, or more importantly, the degree to which exchange did not occur. This knowledge was necessary before considering other pulse programs that rely on transfer of magnetization from the alpha-proton to ^{13}C.

It was critical to obtain an acceptable ^{13}C spectrum of the enzyme, because it would be necessary in later experiments to be able to identify which of the peaks in the ^{13}C spectrum corresponded to tyrosines 48 and 56, the residues adjacent to the valine labels. Therefore, other methods for enhancing the ^{13}C spectrum were examined.

There are three ways in which the quality of the ^{13}C spectrum can be improved. First, the length of time the experiment is run could be extended. This is not possible given instrumental scheduling considerations. The experiment could also be run on a higher field instrument, although this would most likely not enhance the spectrum significantly. Second, the concentration of the sample could be increased. The concentration of the sample is always a concern for the researcher doing NMR on biologically signifcant samples. In this case, the sample of the concentration was approximately 0.5 mM, which is somewhat less than optimal. This sample represents the results of three separate growths on the deuterated medium. Based on considerations for the required time for more growths and the availability of materials, this option was not deemed viable. The third option requires the use of other pulse programs to enhance

sensitivity. A number of different pulse programs are available for the enhancement of insensitive nuclei. Among these pulse programs, the Distortionless Enhancement by Polarization Transfer (DEPT) experiment (Bendall *et al*, 1981) is commonly used, and was chosen as the best method to enhance the ^{13}C spectrum of the doubly-labeled enzyme. The DEPT sequence involves the transfer of magnetization from a carbon-bound proton to the carbon it is bound to. Carbon is then detected directly.

The INEPT pulse program (Morris & Freeman, 1979) would also provide an enhancement to the ^{13}C spectrum, however previous experience in this laboratory (Harpel *et al*, 1995) indicated that the DEPT program yields better results. This is due to the need for absolute accuracy in pulse calibration and coupling constant determination fora successful INEPT experiment.

The DEPT experiment was developed to increase signal strength of nuclei with a low gyromagnetic ratio and/or low natural abundance. The experiment uses a polarization transfer from the sensitive nucleus to the insensitive nucleus to increase signal strength. The enhancement in sensitivity is expressed by the ratio of the sensitive nucleus (from which magnetization is transferred) to the insensitive nucleus (to which the magnetization is transferred), γ_S/γ_I. For a proton to carbon transfer, this represents an enhancement of almost a factor of 4.

The DEPT experiment was prepared such that the polarization transfer will only occur if the sensitive and insensitive nuclei are one bond apart. Therefore, only the carbon nuclei that have attached protons will be observed in the spectrum. This sequence then has the added benefit of being able to "filter" the spectrum such that only the

$^{13}C_{\alpha}$-carbons that have protons attached, in theory only the tyrosine residues, appear in the spectrum. Carbons other than the introduced label will have deuterons rather than protons at the α position and, since the deuteron is a much less sensitive nucleus, these carbons will not be observed.

Figure III-7 shows the ^{13}C DEPT spectrum of the doubly-labeled enzyme. This experiment is expected to show seven peaks, one for each of the labeled tyrosine residues. Examination of the spectrum shows that indeed, seven peaks are present. The large peak at 41 ppm likely represents an overlap of two signals. Although only the C_{α} region of the spectrum is shown in this figure, there are no other peaks present in this spectrum. The observation of only the seven expected peaks indicates that there was a minimal amount of scrambling of the label during incorporation into the growing protein, although it is possible that the DEPT sequence removed any contaminating signals arising from unprotonated carbons from the spectrum.

Another aspect of this spectrum is of great importance: This spectrum shows only seven peaks. If a sample of the isotopically normal enzyme were to be run for a long period of time, all 416 C_{α} peaks would be observed in this region. Moreover, peaks for each of the carbons present in the protein would also appear. It would be impractical to run an NMR experiment for such a long period of time; an experiment of this nature would require many days of uninterrupted instrument time. To avoid this problem, researchers have utilized labeling strategies to incorporate isotopes, such as ^{13}C, into the growing proteins. This approach is similar to the approach taken in this work with one

key difference. Previously, spectra were isotopically enhanced by a uniform labeling strategy. That is, every residue in the protein was labeled with ^{13}C. This labeling strategy is acceptable for smaller proteins and peptides that generate spectra with a manageable number of signals. Enzymes such as PGK are much too large to allow interpretation of spectra from a uniform labeling. To wit: the seven peaks observed in Figure III-7 would be supplanted by 416 peaks. Without expanding this into a third or fourth dimension, analysis of such a spectrum would be a difficult task. The labeling strategy described in this work avoids this pitfall, and produces the simple spectrum shown.

The peaks corresponding to Tyr 48 and Tyr 56 were, however, still unidentified in the carbon spectrum. In order to identify the these two tyrosines, it was necessary to perform an HNCA experiment (Kay *et al*, 1990). Due to instrumental limitations at UTK, another trip to NMRFAM was required. The HNCA experiment was performed on a 500 MHz instrument and a modified version of the traditional HNCA pulse program was used (Volkman, 1996).

The HNCA experiment allows correlation of the amine proton and the C_α of the neighboring residue. In this work, the correlated nuclei are the $^{13}C_\alpha$ of tyrosines 48 and 56, and the amide proton (NH) of valines 49 and 57. Note that this pulse sequence uses an ^{15}N filter. In order for the C_α and amide proton to generate a crosspeak, the magnetization must be transferred, or filtered, through the ^{15}N nucleus the proton is bound to. This is a three dimensional sequence, then, that has been edited to show only

two dimensions in the spectrum. This sequence will also only show a crosspeak if these two residues are adjacent to each other in sequence.

Figure III-8 shows the HNCA spectrum obtained. This spectrum is expected to show two crosspeaks, and two are observed. These two crosspeaks represent the two doubly labeled sites in the protein. This experiment has then identified two out of a possible seven interactions. It is the most important experiment in this work. It shows that by following a directed labeling strategy, it is possible to identify two out of 416 residues of a large protein. It is this experiment that validates the labeling strategy developed for this dissertation.

The advantages of the current method over a uniform labeling approach are clear. If a uniform labeling stratgey was used in this case (uniformly labeling with ^{15}N and ^{13}C), crosspeaks would be observed for each of the 416 residues, and the benefits of this experiment for the study of large proteins would be lost. Even if the protein were labeled with $^{13}C_{\alpha}$ tyrosine, as was done here, and uniformly labeled with ^{15}N, there would still be seven crosspeaks observed, again making analysis more difficult.

The HNCA experiment itself also serves as a "scrambling filter". That is, there is a minimal chance of a contaminating crosspeak in the spectrum. In order for such a crosspeak to arise, the label would have to be metabolically broken down and reincorporated into a residue that is adjacent in sequence to the other label, which would have to be similarly degraded and reincorporated. This is a highly unlikely event.

The next step in the characterization of the doubly labeled enzyme was to correlate the $^{13}C_{\alpha}$ with its attached proton. This was accomplished using the pulse

program INV4TP (Bax *et al*, 1983). This experiment allows correlation of protons and an X nucleus, ^{13}C in this case, via heteronuclear zero and double quantum transitions. The main variable in this experiment is the coupling constant chosen to describe the extent of interaction between the $^{13}C_{\alpha}$ and its attached proton; in this experiment a value of 140 Hz was used. Figure III-9 shows the results of this experiment. Only the region of the spectrum corresponding to the expected $^{13}C_{\alpha}/\alpha H$ interactions is shown in this figure. Seven crosspeaks are expected, and seven strong crosspeaks are observed in this figure. A number of weaker crosspeaks are observed, and these are likely due to the small degree of label scrambling that did occur. The seven crosspeaks corresponding to $^{13}C_{\alpha}$-to-proton interactions are labeled in the figure. Also labeled are the two crosspeaks identified in the HNCA experiment (Figure III-8) as belonging to Tyr 48 and Tyr 56. Another strong peak is seen around 42 ppm in the carbon dimension and 3.5 in the proton dimension. Although this peak is strong, it is not likely that this is due to an $\alpha^{13}C$ interaction because its carbon and the corresponding proton chemical shifts are not in the expected range for such a crosspeak.

The final step in the characterization of this enzyme was to correlate the αH and the side chain protons. This was accomplished using an INV4MLNDTP pulse sequence (Lerner & Bax, 1986). This pulse sequence is a two dimensional experiment that allows correlation of protons and X nucleus via zero and double quantum transitions. This experiment is similar to the INV4TP sequence described above, however the INV4MLNDTP sequence utilizes an MLEV-17 sequence for spin lock. This experiment

is then somewhat akin to a heteronuclear TOCSY experiment. Again, the main variable in this experiment is the coupling constant between the $^{13}C_{\alpha}$ and the side chain protons. A coupling constant of 140 Hz was used in this experiment. This experiment is very sensitive to the choice of coupling constant, as will be discussed below.

Figure III-10 shows the results of the INV4MLNDTP experiment. The same region of the spectrum is shown in this figure as was shown in Figure III-9, for ease of comparison. This figure shows the appearance of two strong crosspeaks in the 3.6-2.9 ppm (in D1) region of the spectrum. These crosspeaks represent interactions between the αH and βH's of the tyrosine residues, and these connectivities are highlighted by horizontal arrows. Another weaker crosspeak is present at 3.3 ppm, but is not visible in this spectrum due to the higher cut in the spectrum required for acceptable presentation of the other data in the spectrum. A deeper cut of this spectrum does show this crosspeak. There are seven tyrosines, and as such it is tempting to expect seven visible crosspeaks in this spectrum: one for each tyrosine αH-to-βH interaction. However, it must be remembered that this experiment is extremely dependent on the value chosen for the coupling constant representing the $^{13}C_{\alpha}$-to-βH interaction. As such, it is improbable that one value will be adequate for all seven residues. This is why only two strong crosspeaks are observed in this spectrum. If it was desired to visualize each of the seven possible crosspeaks, a series of experiments would be required. In the current work, this was not necessary, as the presence of the two strong crosspeaks satisfactorily demonstrates the connectivity between the αH and the βH of the same residue.

It is important to bear in mind that this experiment utilizes an MLEV-17 pulse sequence for spin-lock. As such, it is not possible to examine connectivity of the βH's to the ring protons.

D. Discussion

The major objective of this work was to demonstrate the feasibility of a site specific, non-uniform labeling strategy to be used in combination with perdeuteration of the protein. The labeling strategy developed in this work permits detailed structural studies of large proteins by NMR, which is not practical using traditional NMR methods.

Two isotopically labeled enzymes were prepared in this work. Both were labeled so as to allow identification of two out of the total 416 residues of PGK. The first enzyme was not fully characterized due to instrumental limitations. These limitations were overcome, however, and the valine/tyrosine labeled enzyme was completely characterized.

The many deuterated enzyme preparations carried out in our laboratory indicate that the best source of carbon is deuterated algal cell hydrolysate. Due to the unavailability of this component of the growth medium, two alternate sources of carbon were utilized in these preparations, deuterated glucose and deuterated glycerol. Cultures grown on the latter compound showed markedly better growth although the growth was slow and the protein yield was low. Both of which are consistent with results of previous labeling studies in this laboratory for cultures grown on both deuterated algal cell

hydrolysate and deuterated glucose (Pappu, 1996). Even with the unnatural growth conditions, enough protein was obtained to allow all required NMR studies. This fact supports the idea that a labeling study similar to that described here can be undertaken in any other laboratory interested in performing structural studies on a protein. Of course, the growth conditions for another system will have to be optimized. This can be done using small test cultures, as was done in this study.

Again, the main goal of this work was the development of a labeling methodology to allow identification of selected residues of a protein. As such, the work does not extend to further stuctural studies, although such studies are certainly possible. In fact, the deuterated background of this enzyme allows the study of protein-ligand interactions at much lower ligand-to-enzyme ratios. Using this methodology, the traditional requirement that such interactions be studied under high ligand-to-enzyme ratios, which give fast exchange conditions, is no longer necessary.

The tyrosine-valine labeled enzyme allowed identification of two tyrosine residues, Tyr 48 and Tyr 56. However, it is difficult to determine which crosspeak corresponds to which residue. An assessment of the crystal structure of PGK allows a possible insight into this question. One of the crosspeaks is approximately 5 ppm upfield from the other. This suggests that this residue is located in a more hydrophobic local environment. Close examination of the local environments of each of these residues shows that they are indeed found in very similar environments. As such, it is difficult to absolutely assign either of these crosspeaks.

By identifying one or several residues of a protein, it is possible to use these residues as probes of local structure. The work described above can therefore be applied to a variety of different types of studies, including enzyme-substrate interactions (as described above. Also see Pappu & Serpersu, 1994; Pappu, 1996), protein folding studies, site-directed mutagenesis studies and many other studies were conclusive identification of a probe residue is required.

Another advantage of this methodology is that all the NMR spectra required for identification of a residue, with the exception of the HNCA experiment, can be performed using lower field instruments with no special hardware. In fact, even the HNCA experiment can be performed if an instrument with a third channel is available. Such instruments are becoming more and more commonplace. All the spectra shown in this work were taken on a wide bore 400 MHz instrument except for the HNCA experiment. Therefore application of this methodology is not limited by instrumental considerations.

To summarize the work, perdeuterated enzymes were prepared that contained either two or three protonated labels. These labels were not only protonated but contained either $^{13}C_{\alpha}$ or ^{15}N isotope labels. These labels were correlated and allowed identification of two of the seven tyrosine residues found in the sequence of PGK. This labeling strategy demonstrates the feasibility of this approach for the study of proteins that are large by NMR standards. Moreover, upon optimization, this strategy can be applied to the study of any large protein.

Chapter IV

Expression of Yeast Phosphoglycerate Kinase in *Escherichia coli*

A. Introduction

One drawback of studying proteins and other biological samples using NMR is the large amount of sample required for obtaining acceptable spectra. Sample concentrations on the order of millimolar amounts are optimal, with approximately 0.5 mM being the lower limit. Ikura *et al.* (1990) report that protein concentrations of at least 1 mM are required for high signal-to-noise ratios in two and three dimensional experiments. At concentrations of less than 0.5 mM, spectra are not well resolved and extended, impractically long acquisition times are required. Even at a concentration of 0.5 mM, it is sometimes necessary to make additional experimental considerations, such as the acquisition temperature, in order to achieve meaningful results.

The previous chapters have elucidated a successful isotope-labeling strategy for large proteins. However, the amount of protein isolated from the deuterated and isotope-labeled preparations was substantially less than the amount obtainable from cells grown on isotopically-normal media. Although spectra of good quality were obtained using protein purified from the deuterated growths, a larger amount of protein would have allowed faster acquisition times and possibly a greater range of experiments. In fact, with the exception of the HNCA experiment, it is probable that many of the

experiments requiring a higher field instrument would have been able to have been performed on site.

It has been discussed previously (Chapter III, Results and Discussion), that one benefit of doubly labeling a perdeuterated protein using the scheme described is that the possibility of label scrambling is minimized. In the double labeling scheme, the only possible way to observe contaminating peaks due to label scrambling is to have very specific degradations of the labels, followed by reinsertion at several very specific positions in the primary sequence of the protein. However, when a study is done in which only one type of isotope label is to be incorporated into the growing protein, the effects of label scrambling can become more troublesome. An example of this the reincorporation of added tyrosine label into alanine residues in perdeuterated yeast PGK (Chapter II, Results and Discussion).

One way in which label scrambling can be minimized is through the use of an auxotrophic strain. An auxotrophic strain is a strain that is unable to grow in the absence of a given compound. The auxotrophic quality of many strains is utilized as a selectable marker in various screens. Auxotrophic strains are also of great use in isotope labeling studies of proteins. The choice of a strain of organism that is auxotrophic for the residue being included as a label greatly increases the incorporation of that residue into the protein, since the organism is not able to synthesize the compound by its own means. For example, a tyrosine auxotrophic strain cannot grow without the addition of tyrosine to the growth medium, and therefore the tyrosine added to the medium is almost fully incorporated. The use of such a strain greatly reduces the chances of degradation and/or

scrambling of the added label. The studies performed in Chapters II and III were done without the benefit of an tyrosine auxotropic strain of yeast. This is because such an auxotrophic yeast strain was not available.

It was decided to research a system that would allow both overproduction of yeast PGK and allow the expression of this protein in an auxotrophic strain. By satisfying both of these goals, subsequent isotope labelings would exhibit more efficient label incorporation. A number of expression systems are available for the overproduction of a given protein, however the additional requirement that the protein be expressed in an auxotrophic strain of organism suggested expression in an *Escherichia coli* system. Previous experiments indicated that *E. coli* was able to grow well on deuterated media. Expression of yeast PGK in such this organism would allow not only the possibility of overexpression, but also the selection of a variety of obtainable auxotrophic strains.

Several methods of overexpression of foreign proteins in *E. coli* are currently available. It was decided that the best, most direct way to achieve overexpression of yeast PGK in *E. coli* was to use a plasmid containing the T7 RNA polymerase promoter sequence, a strong promoter commonly used in bacterial systems, and to express the yeast protein in a cell line that expressed this polymerase. The general strategy for the expression of the yeast protein in *E. coli* was to use polymerase chain reaction (PCR) techniques to generate a copy of the yeast *PGK* gene. This gene would then be inserted into the plasmid containing the T7 promoter. This plasmid could then be inserted into any one of a variety of auxotrophic strains of bacteria.

Before this cloning project was undertaken, an extensive review of the literature was made in order to determine if this identical expression had already been accomplished. This review indicated that yeast PGK had never before been expressed in bacteria. The reason for this may be that the expression levels of PGK in yeast are substantial, as are most of the glycolytic enzymes, and therefore the need for such a system was minimal. It should also be pointed out that beyond our laboratory, there have been no studies in which yeast PGK was either isotopically labeled or perdeuterated. Again, this suggests that the need for such an expression system was not great.

B. Experimental Procedures

1. Materials

Yeast strain 20-B12 was utilized in these studies. It contained the multicopy plasmid *pCGY219* and both were provided by Dr. G. Hitzeman. Yeast growth media components were obtained from the following sources: casein amino acids, yeast nitrogen base, yeast extract and tryptone were obtained from Difco, and glucose, adenine sulfate, uracil, myo-insotol, pyridoxine, sodium chloride and thiamine were all purchased from Sigma. SOC media used was provided as part of the TA Cloning kit.

The PCR primers were obtained from Research Genetics and the PCR system was optimized using the optimization kit from Idaho Technologies. The plasmid *pT7-7* was provided by Dr. Cynthia Peterson. The TA cloning kit used in these studies was obtained from Invitrogen, and competent BL21 pLysS cells were obtained from Novagen. IPTG and X-Gal were purchased from was purchased from GibcoBRL. The Wizard minipreps

used in plasmid isolations, Taq polymerase and all restriction enzymes were obtained from Promega. The Qiaex system from Qiagen was used in isolation of the DNA fragments.

2. PCR Amplification of The Yeast *PGK* Gene

Yeast were grown on standard media and the multicopy plasmid *pCGY219* was isolated from yeast strain 20-B12 using standard plasmid isolation protocols (Sambrook *et al.*, 1989) and Promega's Wizard miniprep kits. Optimization of PCR parameters, most notably magnesium concentration, was achieved using the Idaho Technologies Optimization Kit. The program used for PCR thermocycling was: denature at 94° C for one minute, anneal at 55° C for 2 minutes and elongate at 72° C for 2 minutes. 30 to 50 cycles were used.

The primers used in PCR amplification of the gene introduced a new *Nde*I restriction site at the 3' end of the gene. This was necessary to allow a subsequent in-frame ligation of the amplified gene. PCR was then performed on six different samples, each containing different concentrations of magnesium, template and primer. The PCR products were then analyzed on 1 % agarose gels to determine which reactions gave products that were of the correct molecular weight, and contained no contaminating side-products of the amplification process.

3. Ligation of The Fragment Into *pT7-7*

After a suitable PCR product was identified, based on analysis of an agarose gel, the fragment was then ligated in the TA cloning vector *pCRII* using the methodology of the TA cloning kit. Three different ligations were performed, each containing a different concentration of PCR product. "One Shot" competent cells were transformed with the ligation product and spread on plates containing kanamycin that had already been treated with X-Gal. The plates were spread with either 50 or 200 μL of transformed cells.

A total of six white colonies were then picked from each of the plates that showed bacterial growth, giving a total of 36 colonies analyzed. These samples were then grown in LB medium containing ampicillin (50 μg/mL), and the plasmid isolated. The plasmid was examined by restriction digestion using the *Eco*RI cut site located on *pCRII* on either side of the inserted gene to determine the presence of the yeast *PGK* gene.

The orientation of the insert was then examined using a series of restriction digests. The plasmid was digested with *Kpn*I to determine orientation. This digest was accompanied by a *Cla*I/*Not*I digest to further confirm the orientation of the insert. The presence of the insert was reconfirmed by a *Bam*HI/*Xho*I digest.

Plasmids identified as those containing the yeast *PGK* gene in the proper orientation were then digested with *Nde*I and *Bam*HI. This digest was used to confirm the presence of the *Nde*I restriction site introduced by PCR and to be certain of the integrity of the *Bam*HI restriction site downstream of the *PGK* gene that would be used in the isolation of the *PGK* gene from the plasmid.

Upon confirmation of the presence of both of these restriction sites, the same digest was redone and the digested fragment was isolated from the agarose gel using the Qiaex system from Qiagen. The expression plasmid, *pT7-7*, was also digested with the enzymes *Nde*I and *Bam*HI. The purified PGK fragment was then ligated overnight into the digested *pT7-7* vector. Competent DH5α cells were then transformed with *pT7-7* containing the yeast *PGK* gene. These cells were then grown overnight at 37° C in the presence of ampicillin (50 μg/mL) on agar plates. On the plates that exhibited growth, approximately 12 colonies were selected and grown up in LB/ampicillin media overnight. The transformed plasmid of the cells of these cultures was isolated and a restriction digest performed. In this analytical digest, the plasmid was again digested with *Nde*I and *Bam*HI to yield the 1.3 kb fragment corresponding to successful ligation of the yeast *PGK* gene. The gene was sequenced to be certain that there were no mutations introduced by the PCR reaction and that the entire gene was present. Competent BL21(DE3)pLysS cells were then transformed with the plasmid known to contain the PGK insert. For cultures that were confirmed to be carrying the yeast *PGK* gene, permanent stocks were made and stored in 10 % glycerol at -80° C.

4. Analysis of Yeast PGK Production by Transformed BL21(DE3)pLysS Cells

Initial analysis of the induction of yeast PGK was done as follows: Cell cultures that were known to carry the yeast *PGK* gene were grown overnight in LB media containing 100 μg/mL ampicillin. A 50 μL aliquot of these cells was then added to 5 mL LB/Ampicillin media and grown for 2 hours. At 2 hours, one mL was removed as an

uninduced control. IPTG was then added to each culture to a final concentration of 1 mM. Cultures were then grown overnight at 37° C. The following morning, a 1 mL aliquot of cells was removed and pelleted by centrifugation. This sample was then resuspended in 100 μL of SDS gel loading buffer and incubated at 95° C for three minutes. 40 μL of the crude extract was run on a 10 % SDS protein gel.

To determine the solubility of the yeast protein, a similar procedure was used (Gibson, 1997). Cells were grown and harvested as above, however a different method of lysis was used. Cells were pelleted, and the residual media removed by aspiration. The cells were resuspended in 25 μL of 0.1 M Tris-HCl, pH 8.0. The cells were then incubated with lysozyme (0.375 mg) and DNAse (0.04 mg) for 15 minutes at 37° C for 15 minutes. They were then vortexed, frozen in dry ice/ethanol and then incubated at 37° C for 15 minutes. This procedure was performed three times. The samples were then centrifuged in a microfuge at top speed for 15 minutes. The supernatant was removed and mixed with an equal volume of 2X SDS gel loading buffer. The pellet was mixed with 25 mL of 8 M urea and 25 mL of 2X SDS gel loading buffer. The samples were boiled for 5 minutes and loaded on a 10 % protein gel. Both the soluble and insoluble fractions were analyzed by SDS PAGE for the presence of yeast PGK.

The optimal induction time for the bacterial system was determined by inoculating 5 mL of LB/Ampicillin (50 μg/mL) media with overnight cultures also grown on LB/Ampicillin media. These cultures were grown at 37° C for two hours. A one mL aliquot of cells was removed and the cells pelleted. IPTG was added to a final concentration of 1 mM. One mL aliquots of cells were removed after 1,2 3 and 12 hours.

Aliquots taken at 1,2 and 3 hours were stored at 4° C until the 12 hour aliquot was harvested. Cells were lysed using the freeze/thaw procedure described above, and the soluble fraction was run on a 10 % protein gel.

No difference in expressed PGK levels was observed between 3 hours and overnight inductions and therefore 3 hours was used as the optimal induction time for subsequent experiments.

Comparative expression levels of PGK in the bacterial system and expression in the yeast system were made by comparing the yield of protein from cultures of identical volumes.

C. Results

The first step in the expression of yeast PGK in an *E. coli* system was the generation of a copy of the gene from the yeast multicopy plasmid *pCGY219*. Very little is known about the location of restriction sites on this plasmid, due to the minimal notes provided by Dr. Hitzeman. Unique cut sites were identified on upstream and downstream sides of the gene, but the use of these sites would yield a fragment that also contained a substantial number of bases upstream of the gene of interest. To avoid this problem, as well as the possibility of an out-of-frame ligation, it was decided to introduce an *Nde*I restriction site at the beginning of the *PGK* gene. This was accomplished through the use of PCR. By utilizing a primer that included the *Nde*I site, the PCR product would be assured of containing the new restriction site. The introduction of this

site involved a one-base mutation in the sequence of PGK. Primers used in PCR are shown in Figure IV-1.

The conditions of the PCR reaction were optimized for amount of template, amount of primer and concentration of magnesium. It was determined that the optimal conditions for PCR were 10 pM template, 1 pM primer and 1.5 mM magnesium. These conditions were used in all subsequent PCR reactions, although the reaction volume was increased to 25 μL and the components scaled up accordingly. The PCR products were analyzed on a 1 % agarose gel and the results of the reactions are shown in Figure IV-2. The sample were determined to have the correct size reaction product and this fragment was then ligated into the TA cloning vector *pCRII*. Following the overnight ligation of the fragments, competent "One Shot" cells were transformed with the vector. These transformed cells were then spread on LB/Kanamycin/X-Gal plates and grown overnight at 37° C.

The *pCRII* vector allows for blue-white screening; white colonies indicate the presence of the insert, while blue colonies indicate lack of the insert. Six white colonies from each plate showing growth were selected for analysis, giving a total of 36 samples analyzed. These colonies were then grown up in LB-Ampicillin media. The *pCRII* vector was isolated from the cells and analyzed by restriction digest to confirm the presence of the yeast PGK insert. This digest utilized the *Eco*RI sites located on *pCRII*. These sites flank the insert on *pCRII* and were therefore identified as a good way to determine the presence or absence of the insert. These digests were run on a 1 % agarose gel and the results are shown in Figure IV-3. The *PGK* gene consists of

approximately 1.3 kb and was found to be present in the majority of the ligation products. It was decided to use the vectors L6502, L42002 and L62002 to confirm the presence and orientation of the insert.

The orientation of the insert was determined through the use of a series of double digests. Three different digests were performed. *Kpn*I and *Cla*I/*Not*I digests were used to determine the orientation of the insert. *Bam*HI/*Xho*I digests were used to reconfirm the presence of the insert. The results of this experiment confirm that the insert is present and are shown in Figure IV-4. In the *Kpn*I digests, if the gene is inserted in the correct orientation, the expected fragments are of lengths 750 and 4400 bp, which is what is observed. Similarly, in the *Cla*I/*Not*I digests, if the gene is inserted in the correct orientation, the expected fragments are of lengths 1260 and 3885 bp, and these fragments are observed on the gel. The insert is clearly present as proven by the presence of the expected 1300 and 3900 bp fragments generated from the *Bam*HI/*Xho*I digest as shown on the gel. Based on this gel, it was decided that the plasmids to analyze further and potentially use in later subcloning work were L42002, L6502 and L62002.

The next step in the analysis of these plasmids was to confirm the presence of the introduced *Nde*I restriction site. It was also necessary to confirm the presence of the *Bam*HI site, since both of these sites were to be used in the excision of the *PGK* gene for later subcloning. Figure IV-5 shows the results of the *Nde*I/*Bam*HI digest. This gel shows that both the *Nde*I and *Bam*HI sites are present, as evidenced by the presence of a 1.3 kb fragment corresponding to the PGK insert.

The cultures containing the the vectors showing favorable results were then grown again in LB/Ampicillin media and the *yPGK/pCRII* plasmid isolated. The identical digest was done (using *Nde*I and *Bam*HI) and again separated on a 1 % agarose gel. The fragment corresponding to yeast PGK was then excised from the gel and isolated from the agarose using the Qiaex system. This treatment completely solublized the fragment. Similarly, the expression vector *pT7-7* was digested with the same two restriction enzymes, removed from the gel and isolated using the Qiaex system. The yeast PGK fragment was then ligated into the expression vector *pT7-7*. This vector was chosen because it contained the strong T7 polymerase promoter and convenient restriction sites in the multiple cloning region for the insertion of the PGK fragment. A plasmid map of the ligation product is shown in Figure IV-6.

Competent DH5α cells were then transformed with the ligation product. The transformed cells were spread on LB/Ampicillin plates and grown at 37° C. Colonies were then selected and analyzed for the presence of the gene of interest using restriction digests. The plasmid was first isolated and digested with *Nde*I and *Bam*HI. The results of this digest are shown in Figure IV-7. This experiment demonstrated that the gene was, in fact, present in the *pT7-7* expression vector. It was not necessary to determine the orientation of the insert in this vector because the ligation was a "one-way" ligation that took advantage of two unique restriction sites.

Before proceeding with induction studies it was necessary to sequence the *PGK* gene to be certain that there were no random mutations introduced by the action of Taq polymerase. Two samples, L42002 and L6502, were sequenced and it was determined

that (1) there were no mutations present, and (2) that the *Nde*I restriction site was present (sequencing gel not shown). Upon confirmation of the correct gene sequence, the expression vector *pT7-7* containing the sequenced PGK insert was transformed into the cell line that was to be used for expression, BL21(DE3)pLysS. The following experiments all involve this expression system: the insert-containing vectors L6502 and L42002 in the BL21(DE3)pLysS cell line. The insert-containing vector L62002 was not used in subsequent steps since it had not been sequenced. These expression systems will hereafter be referred to simply as "L6502" and "L42002".

Upon confirmation of the integrity of the *PGK* gene, it was necessary to determine whether the protein was being overexpressed, whether it was soluble, and what the optimal conditions for maximal protein induction were. The first experiment to be performed was designed to test whether or not the system was overexpressing the yeast gene. LB/Ampicillin media were innoculated with L6502 and L42002 cultures (grown overnight) and grown for two hours. Uninduced control samples were removed. The *T7 polymerase* gene was then induced by the addition of IPTG, thereby producing yeast PGK. Cells were grown at 37° and harvested after an overnight growth. Cells were lysed by boiling in SDS gel buffer and run on SDS-PAGE. The protein gel for each of the systems (L6502 and L42002) is shown in Figure IV-8. This gel shows both the induced and uninduced controls. A comparison of the induced and uninduced samples on this gel showed that the bacterial system was, in fact, overproducing the yeast PGK protein.

The next step was to determine if the overexpressed yeast protein was expressed in soluble form. Cells were grown and harvested as above, however in this experiment,

the cells were lysed by a combination of three freeze/thaw cycles and the action of lysozyme and DNAse, as decribed above (Chapter IV, Experimental Procedures). Following lysis, both the soluble and insoluble fractions were analyzed by SDS PAGE. The results of this experiment are shown in Figure IV-9. From this data it is clear that the protein is being expressed at high levels and in soluble form.

The next step was to determine the optimal time for harvesting the cells following induction. Again, following innoculation of LB/Ampicillin media with overnight cultures of L6502 and L42002 and removal of an uninduced sample, the cells were induced with IPTG and harvested at 1,2 3 and 12 hour time points. The cells were lysed as described above and the soluble fraction run on a 10 % protein gel. Figure IV-10 shows the results of this experiment. The expression level of yeast PGK clearly increases from 0 to 3 hours, however there is not a large difference between the 3 hour time point and the 12 hour time point. It was decided that, since there is not a great difference in the level of expressed protein, a 3 hour induction time would be used. The 3 hour induction time was selected to minimize the possibility of degradation of the overexpressed protein, and was used in all subsequent experiments.

An assessment was then made of the yield of yeast PGK from both the *E. coli* expression system and the present yeast expression system. A number of preparations were done on the bacterial system. It was determined that the enzyme was being expressed in active form, however, in each preparation, the yield of protein obtained was either similar to or less that that obtainable from the yeast expression system already in place in our laboratory. The determination of the amount of protein produced was made

based on activity assays and SDS-PAGE analysis. In an effort to increase the yield of protein from these preparations, a number of variations to the purification protocol were made. These alterations included omission of column chromatography steps, omission of pH precipitation steps, and omission of a protamine sulfate precipitation step. Unofortunately, none of these modifications enhanced protein yields, which were routinely 4 mg or less per 50 mL culture. In order to be certain that the low protein yield was not due to the inability of the cells to replicate, growth was monitored. Growth curves from these monitored growths showed that the cells were growing to normal levels and exhibiting normal growth profiles. Based on these results, further characterization of the bacterially-expressed protein was not pursued.

It was also observed that the expression levels of yeast PGK decreased over time. That is, the same frozen stock of cells did not produce the same amount of yeast PGK in every growth. This effect has been observed previously and reported to be not uncommon. In order to maintain the high expression levels of PGK, it would be necessary to retransform the cells before each growth. Had the protein yield warranted it, retransformation of cells prior to each growth would have been incorporated into the growth/purification protocol.

D. Discussion

The results of this series of experiments show that it is possible to express yeast PGK in soluble form and at high levels in an *E. coli* expression system. However, the

drawback of this expression system is that the yield of PGK was not significantly higher than that achievable from the current yeast system. Based on this result it was decided that the yeast expression system remained the best method for expression of deuterated and isotopically-labeled PGK.

Despite the low levels of protein produced, the two major goals of this series of experiments were met. Firstly, it was demonstrated that it was possible to express yeast PGK in a bacterial system, and that the protein was expressed in soluble form. Prior to these experiments, this had never been accomplished.

Secondly, this work generated a plasmid that can be transformed into auxotrophic strains of bacteria to enhance the isotope-labeling efficiency of yeast PGK . Although the expression level of the protein in the bacterial system is not substantially greater than that achievable in the yeast system, the possibility of protein production is present. This is particularly beneficial when it is desired to label a residue such as glutamic acid, which is extremely susceptible to label scrambling due to its central role in a number of degradation pathways. It is possible that in a case such as this, the bacterial expression system could be effectively employed.

Moreover, purification of the labeled enzyme from the bacterial system is not expected to pose insurmountable obstacles. The bacterial cells are easily lysed using the same protocol used in the yeast PGK isolation procedure (data not shown), and the existing purification protocol for the yeast enzyme can be used without modification. The only possible problem with isolation of the yeast enzyme from the bacterial system could arise from the presence of the bacterial form of PGK. However, it is possible to

separate these two enzymes on the basis of their differences in charge. Fifis & Scopes (1978) report that while the yeast enzyme is adsorbed on CM-cellulose at pH 5.8, the *E. coli* enzyme is not. For purification of the *E. coli* enzyme, adsorption on DEAE-cellulose at pH 8.2 was required, followed by gradient elution. Therefore, although extensive research was not performed on the separation of these enzymes, this study suggests that, given sufficient optimization of conditions, such a separation is plausible.

CHAPTER V

Small Angle Neutron Scattering Studies of Phosphoglycerate Kinase and Glyceraldehyde-3-Phosphate Dehydrogenase

A. Introduction

When characterizing the physical properties of a protein or enzyme, researchers usually prefer to work with the sample in dilute solutions. These dilute solutions provide a sample environment that reduces the possibility of non-ideal behavior due to interactions between particles in solution, thereby greatly simplifying the analysis. Some of the complications leading to non-ideality include solute-solute interactions and interaction of the solute with salts, buffers and other components of the solution. When performing enzyme activity assays, protein concentrations of only 10^{-9} to 10^{-6} M are required. These concentrations are significantly below 10^{-2} M, the concentration at which proteins become substantially aggregated (Mastro & Keith, 1984). As such, extensive work has not been done on the properties of enzymes and substrates in concentrated protein solutions.

It is interesting that there has been only limited work on the properties of enzymes in concentrated protein solutions, especially in light of the concentration of proteins typically found in the cell, which ranges from 250-400 mg/mL (Srere, 1987). This

concentration is much higher than the concentrations of proteins for which nonideal behavior has already been identified. Of this cellular concentration, a substantial portion is attributable to the glycolytic enzymes. Therefore, it is not unreasonable to expect that the glycolytic enzymes interact closely with one another.

There are two broad groups of enzyme-enzyme complexes: isolatable stable multienzyme complexes, and those complexes associated by weak interactions, that can be classified based on their equilibrium constants. The former group has been extensively studied and examples include fatty acid synthase and the pyruvate dehydrogenase complex. The latter group is less studied and an example of this type of interaction is the subject of this chapter.

A number of instances of weak enzyme-enzyme interactions involve enzymes that are sequentially related in metabolic pathways (Srivasta & Bernhard, 1987). The issue raised by this observation centers on the possible advantages the cell may achieve via these associations. Such associations can greatly affect the kinetics and thermodynamics of catalyzed reactions (Keleti & Ovadi, 1988). The occurrence of these weakly associating complexes has been demonstrated for a number of enzyme pairs, including glyceraldehyde-3-phosphate dehydrogenase and lactate dehydrogenase (Srivasta & Bernhard, 1985), aldolase and fructose 1,6-diphosphate (Horecker *et al.*, 1981) and citrate synthase and malate dehydrogenase (Datta *et al.*, 1985).

The exact role that weak interactions between soluble enzymes play is not clear. However, it has been proposed that these interactions may play a role in metabolism by either facilitating channeling of products from one enzyme to another, or by the induction

of a conformational change in one or both enzymes of the pair, leading to altered catalytic or regulatory properties (Clarke & Masters, 1975; Keleti *et al.*, 1977). The idea of channeling is particularly suited to enzymes that are sequential in a given pathway, such as glycolysis and the citric acid cycle. The enzyme pairs cited above are all purported to transfer their common metabolite by a direct transfer, or channeling, mechanism and function as part of a larger pathway. In fact, the idea of channeling common metabolites between enzymes has been put forth as one of the foremost mechanisms of regulation of several pathways (Keleti & Ovadi, 1988; Freidrich & Hajdu, 1986).

Although the association of sequential enzymes and the channeling of their common metabolites is well documented, there is still some disagreement as to the validity of this model. The vast majority of evidence supporting direct transfer is based on kinetic studies. It has been argued that the assumptions made in these studies were erroneous (Vas & Batke, 1990; Kvassman & Pettersson, 1989). It has also been argued that there has been no successful attempt to detect such a complex by physicochemical methods (Vas & Batke, 1981). It may be that these transient complexes are difficult to detect using traditional biophysical methodology.

One such enzyme pair is of particular interest to our laboratory, the enzymes glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoglycerate kinase (PGK). These enzymes are sequential in the glycolytic pathway.

The reactions are linked as follows:

$$\text{3-PGA} + \text{ATP} \overset{\text{PGK}}{\Leftrightarrow} \text{ADP} + \text{1,3-DPG} + \text{NADH} \overset{\text{GAPDH}}{\Leftrightarrow} \text{G-3-P} + \text{NAD}^+ + \text{P}_i$$

As discussed in previous chapters, PGK is one of several enzymes that constitute the main focus of our studies. A number of other laboratories have studied the interaction of this same enzyme pair. Foremost among the studies of these enzymes is the kinetic study of Weber and Bernhard (1982) on the halibut muscle enzymes. In this kinetic study, evidence is presented that argues for the existence of a complex between the enzymes. Weber and Bernhard monitored the transfer of 1,3-diphosphoglycerate through the use of the coupled enzyme assay used to assay PGK activity described in this dissertation. They describe the direct transfer using a Michaelis-Menten scheme where PGK•1,3-diphosphoglycerate acts as the substrate for GAPDH. Such a complex has the advantage of vectorial regulation, as discussed above, and also explains the lack of any detectable 1,3-diphosphoglycerate in solution (Bloch *et al.*, 1971). The Weber and Bernhard study presents evidence for a direct transfer of the common metabolite 1,3-diphosphoglycerate between the two enzymes. A model is presented to explain the transfer mechanism. The Weber and Bernhard study was the first study to clearly demonstrate such a transfer.

A number of other studies have supported the results of Weber and Bernhard. A study on the rabbit muscle enzymes gave physical evidence in the form of electron micrographs of the association (Sukhodolets *et al.*, 1988). Another approach that has been taken is to immobilize GAPDH on a Sepharose column. In the study by Ashmarina

et al. (1984), GAPDH was covalently attached to a cyanogen bromide activated column. PGK was then passed over the column and was observed to bind to the immobilized GAPDH. It was found in this study that three PGK molecules were bound per one GAPDH tetramer. The tetrameric form is the naturally occuring form of GAPDH. The *E. coli* enzymes were studied in the same way by Khoroshilova *et al.* (1992). This group immobilized GAPDH on Sepharose and found that a complex with PGK was formed having a stoichiometry of 1.77 PGK molecules per tetramer of GAPDH. This non-integral stoichiometric ratio may be a product of either weak or multiple equilibria.

The association of the enzymes has also been studied from a theoretical computational approach (Warwicker, 1989). In this study, a method for calculating the electrostatic potential energy between two molecules is presented. Once the potential energy is calculated, the interaction energy is then calculated. This these energies are then layered over the crystal structures of the two molecules and the probability for interaction is then assessed. This probability is based on the number and location of regions on each molecule that are appear to be favorable interacting regions, typically containing large numbers of negatively charged residues. Essentially, this study is a modified docking study. Warwicker found that, when he applied his algorithm to the case of PGK and GAPDH, it showed that intermolecular interactions can persist over large surface regions of both enzymes. Moreover, a possible substrate transfer configuration emerged. This configuration was composed of two PGK molecules binding on opposite sides of one GAPDH, with each PGK binding along the

substrate-binding groove on GAPDH. It should be noted that in this study the tetrameric form of GAPDH and the open conformation of PGK were assumed.

The above literature strongly suggests the possibility of interaction between PGK and GAPDH. However, the majority of the evidence favoring interaction is kinetically-based and, as was pointed out by Vas and Batke (1981), there has been no physicochemical evidence detailing this interaction, nor has the complex itself been isolated. Moreover, although the immobilized GAPDH studies yield convincing data, it can be argued that the non-physiological nature of these experiments require a close assessment of their applicability. In another fluorescence study by Sukhodolets *et al.* (1989), it was discovered that although the FITC-labeled PGK did appear to associate with GAPDH, the fluorescent dye itself may be mediating the interaction. This study is another example of the problems associated with studying comparatively weakly interacting complexes.

In an effort to obtain further biophysical evidence of the PGK-GAPDH interation in solution, a small-angle neutron scattering (SANS) study was undertaken. SANS was chosen as the primary technique in this study for its ability to give information about the shape and geometry of the complex, as well as its ability to give information about the stoichiometry of the interaction. Moreover, this techinique is non-destructive to the sample, which allows experiments requiring a long period of time to be carried out without fear of sample degradation. Most importantly, SANS is a solution-based technique. This is particulary advantageous for studying interactions between proteins

known to exist in soluble form in the cell. The data from an experiment such as this is much more representative of the natural cellular conditions.

In order to appreciate the results of this study, some knowledge of SANS is required. To that end, a short introduction to the technique, as well as a brief description of the instrumental apparatus is presented in the following paragraphs. The SANS camera begins with a neutron source, such as a nuclear reactor. The neutrons are monochromated to a wavelength useful for material science (typically 1-10 Å), and collimated by apertures and waveguides to produce a beam of desired cross section and divergence. This beam passes through the sample. Most of the beam is either absorbed or passes through the sample unchanged. A small portion of the neutrons is scattered by the sample and is counted on a two dimensional detector as a function of scattering angle. The observed scattered distribution is caused by the scattering density fluctuations in the sample.

The angular dependence of the scattering intensity from a solution of particles can be written

$$I(q) = nS(q)P(q) \qquad \text{Eqn V-1}$$

where n is the number density of the particles, P(q) is the particle form factor, and S(q) is the solution structure factor. The form factor describes the scattering arising from the particle itself, and the structure factor describes the scattering from the arrangement of these particles in solution. The scattered angle is expressed by the parameter q, where $q = (4\pi/\lambda)(\sin 2\theta)$, and 2θ is the scattering angle for the neutron.

The radius of gyration is a concept that describes the spatial extent of some volume in terms of a single number for a homogeneous particle or, additionally, the relevant location of different regions in a heterogeneous particle. It is defined by the integral:

$$R_G^2 = \int \rho(r^2)\, dV / \int \rho\, dV \qquad \text{Eqn V-2}$$

This can be analytically evaluated for many homogeneous geometric shapes, such as ellipsoids, cylinders, etc. R_G is usually measured by approximating the low q part of the

scattering curve (assuming S(q) ~1) by a Gaussian function:

$$I(q) = I(0) \exp(-[q\,(R_G/3)]^2) \qquad \text{Eqn V-3}$$

Therefore, a plot of ln[I(q)] vs q^2 will give two independent particle parameters, R_G and I(0), which together, provide powerful constraints on modeling systems of interacting particles. I(0) is the scattering intensity at q = 0, and is proportional to the square of the molecular weight of the scattering particle (in the case of a single particle type). A model must have been successfully screened for consistency to the experimental ellipsoids (to represent the proteins) or the coordinate files from the crystal structures, or a combination of both, assuming S(q) = 1, to be accepted.

One final property of the sample that needs to be addressed is the issue of the degree of scattering from a species. The scattered intensity from a particluar protein is a strong function of its isotopic composition. Each atom coherently scatters neutrons proportional to b, its neutron scattering length. The scattering lengths of hydrogen and deuterium are very different. The scattered intensity of the whole protein is proportional

to the square of its contrast with the solvent, where the contrast is the difference in scattering length densities:

$$[\Sigma (b/V)_{particle} - \Sigma (b/V)_{solvent}]$$

where V is the particle volume and the sums are over either all the protein nuclei or solvent nuclei in this volume. This difference in b-factors is one of the main advantages of SANS studies over small-angle x-ray scattering (SAXS). The most common solvent for protein work is water, which has a scattering length density of -0.0562 cm^{-2}. However, the composition of the solvent can be varied by the substitution of a fraction of water for deuterium oxide. Deuterium oxide has a scattering length density of 0.6404 cm^{-2}. Therefore, the scattering from the solvent itself can be varied by varying the ratio of H_2O to D_2O. As the concentration of D_2O is increased, the most of the exchangable protons will exchange with the deuterated, which has a negligible effect on the solvent for relatively dilute solutions. By increasing the concentration of D_2O, the scattering of the solvent approaches the scattering from the sample itself. The point at which the scattering from the solvent is equal to the scattering from the sample is known as the contrast matching scattering density. At this concentration of D_2O, 41.5% D_2O for proteins, the scattering from the sample is indistinguishable from the solvent and the sample is effectively silenced from the scattering profile. This technique is known as contrast matching.

Contrast matching has the potential to give substantial advantages for the study of protein-protein interations. The technique can be used to cloak one of a pair of proteins

and allow study of the structure of the other protein, but to do this is not a trivial matter. In order to contrast match out a protein, the protein must be composed, to a high degree, of a material having a different scattering length density than that of typical proteins. This can be accomplished by preparing the protein in a fully deuterated form. As discussed in earlier chapters, we have been able to prepare PGK in its fully deuterated form in our laboratory.

The general strategy for the study of the interaction between PGK and GAPDH in solution was as follows: A titration of protonated PGK and protonated GAPDH was performed and SANS was used to study six equilibrium mixtures. Attempts to model this data were not successful, either because of the presence of multiple equilibria or significant deviations of S(q) from unity. Therefore, mixtures of deuterated PGK and protonated GAPDH in 41.5% deuterium oxide were studied in an effort to locate the position, number and equilibrium constant for the association of the PGK with the large GAPDH molecule.

This challenging work represents a novel approach to the study of the proposed interaction between these two glycolytic enzymes. It is one of the few attempts to use contrast matching techniques to study protein-protein interactions using a large fully deuterated protein. Moreover, these experiments have the potential to provide the first physical evidence of these proteins interacting in solution, giving insight into the structural nature of the complex.

B. Experimental Procedures

1. Materials

Phosphoglycerate kinase used in this study was isolated from *Saccharomyces cerevisiae* strain 20B-12 containing the multicopy plasmid *pCGY219*. Yeast nitrogen base and casein amino acids were obtained from Difco; glucose adenine, uracil, myo-inositol, pyridoxine and thiamine were obtained from Sigma. Deuterated glycerol, deuterated glucose, deuterated acetic acid and deuterated ethanol were obtained from MSD Isotopes and Isotech. Yeast glyceraldehyde phosphate dehydrogenase was obtained from Sigma.

2. Yeast Growth on Protonated And Deuterated Media

Protonated PGK was grown as described above (Chapter II, Experimental Procedures) using glucose as the carbon source.

Fully deuterated PGK was grown using the methods described above (Chapter II, Experimental Procedures). Three different preparations were used, and the carbon source was deuterated glucose for the first growth and deuterated glycerol for the second two growths. Cells were grown to late log phase and harvested by centrifugation. Cells were stored at -80° C until used.

The deuterated enzyme was isolated as described in Chapter II, Experimental Procedures, following the method of Kuntz and Kreitsch (1982) with slight modifications. The isolated enzyme was stored at -20° C until used in experiments.

3. Small-Angle Neutron Scattering (SANS) Measurements

Titration experiments of protonated PGK with protonated GAPDH were performed on the 30m SANS camera at ORNL. Concentration series of protonated PGK alone and protonated GAPDH alone were also measured. Plotting c/I(0) for the pure solutions gives the second virial coefficient for these pure systems, which is the simplest way to model the deviation of the solutions from ideality. For ideal solutions, c/I(0) is constant, but these solutions of pure protein have a non-zero slope for c/I(0) versus c. The samples were in deuterium oxide of pathlength 5 mm. The data were reduced with the program SPECTOR (author John Hayter).

A contrast match series ($[I(0)]^2$ versus percent deuterium oxide) was measured for protonated GAPDH, also at ORNL. The cell pathlength varied from 1 mm to 5 mm. The standard textbook match point for proteins (42% deuterium oxide) was effectively observed (41.5% ± 2% deuterium oxide).

Due to subsequent inaccessibility to the SANS camera at ORNL, it was necessary to continue data collection elsewhere. The NIST Facility in Gaithersburg, MD was the most accessible for these studies. Two visits to NIST were made to study samples containing deuterated PGK and protonated GAPDH in 41.5% deuterium oxide. The cell pathlength used in these experiments was 1 mm. The concentration of GAPDH was about 25 mg/mL in both sets of experiments and the concentration of dPGK was about 4.9 mg/mL in the cuvette for the first set of experiments and 12 mg/mL in the cell in the second set of experiments.

The buffer used in these studies was 100 mM NaCl/ 50 mM Tris-HCl, pH 7.5. Buffers were prepared in H_2O, D_2O and 41.5% D_2O. It is known that a 41.5% D_2O solution accurate reflects scattering from a protonated protein sample, and is accepted as a standard. All solutions were centrifuged before use. Purity of the dPGK was determined by SDS-PAGE. Solutions prepared for contrast matching experiments included Tris-NaCl buffer prepared in 100% D_2O and a solution of 2 mg/mL GAPDH in 100% D_2O buffer. These solutions were prepared to calibrate the intensity of the NIST source. Other solutions prepared included 41.5% D_2O Tris-NaCl buffer to calibrate the solution scattering, and stock solutions of GAPDH in 41.5% D_2O Tris-NaCl buffer and dPGK in 41.5% Tris-NaCl buffer.

4. Isolation of Protonated And Deuterated PGK

Isolation of protonated PGK was achieved using essentially the method of Kuntz and Kreitsch (1982).

Three different deuterated growths were required for this work. The first growth utilized deuterated glucose as a carbon source. The composition of the medium was the same as previously described. Growth of these cells was typically slow, and a wet weight of approximately 18 g of cells was obtained. Isolation of the enzyme was, again, as previously described. The yield of enzyme from this prep was about 3.25 mg, which was determined to be sufficient for scattering studies.

The second growth was performed using deuterated glycerol as the carbon source. As before, two growths were pooled for this growth. Again, the growth was slow and a wet weight of approximately 14 g of cells was obtained. Isolation of the enzyme yielded about 5 mg of pure protein.

The third growth also utilized deuterated glycerol as the carbon source. 4 mg of enzyme was isolated from this growth.

C. Results And Discussion

1. SANS Titration Studies at ORNL

The detection of the PGK:GAPDH complex was first attempted by titrating protonated GAPDH with protonated PGK. Six different concentration ratios were studied and are summarized in Table V-1. The R_G's of the mixtures varied approximately with the values of R_G's for the pure proteins, implying that the dissociation constant for the complex was high, and that modeling this system would be numerically challenging. The mixture R_G's and I(0)'s were somewhat higher than those expected for the sum of the two pure protein values, supporting the conjecture that the proteins do, in fact, associate in solution. However, attempts to model the system on the basis of a single complex species formed (i.e. GAPDH•PGK_X where X was varied from 1 to 4, since GAPDH exists as a tetramer), were not successful. There are several possible reasons for this: (a) the salt concentration in the solution was too low to adequately suppress the S(q) factor, (b) two or more complexes were forming, or (c) small errors

were introduced in the preparation of the solutions used, which were amplified by the nature of the data analysis.

2. Contrast Matching Studies at The NIST Facility

Deuterated PGK had been previously prepared in anticipation of these experiments and had been stored frozen at -20° C. It was decided to use this sample in the contrast matching studies. It had been previously determined that a buffer composed of 41.5% D_2O would effectively equate the scattering of the solvent to that of the GAPDH. Because GAPDH is much more massive than PGK, and scattering intensity varies with the square of the molecular weight, the analysis was expected to be numerically more manageable by matching out the larger species. All buffers and proteins were centrifuged for 10 minutes in a microfuge prior to their use.

Scattering from dPGK was studied first. The R_G of this protein was determined to be 23.9 ± 1.1 Å. The scattering profile is shown in Figure V-1. This value was in good agreement with the R_G for isotopically normal PGK determined earlier, which was 24.0 ± 0.3 Å (Henderson *et al.*, 1994). However, the scattering intensity was decreased by a factor of 10, from 0.43 observed in previous experiments to 0.028. This suggested that the enzyme was aggregated and inactive. Although centrifugation did serve to remove the large aggregates, the smaller aggregates remained. The presence of these small agreagates is evident from the Guinier plot, which shows aggregation at low Q values. Further interpretation of the results from this experiment is not possible due to this

sample degradation. Therefore, it was decided to attempt the experiment again at a later date.

For the second attempt at these experiments, fresh deuterated PGK was prepared, although only about 4 mg of enzyme was obtained. The strategy for these experiments was the same as for the previous attempt, and identical fresh buffer solutions were prepared.

Required solutions were prepared the day before the experiments were to be performed. Again, solutions were centrifuged to remove any aggregates. The buffers and protein samples were sent overnight to NIST to minimize the time they would be unrefrigerated and to minimize the possibility of aggregation. Despite these precautions, the dPGK was found to aggregate extensively and the experiments were not able to be performed as anticipated.

Initial titration studies indicate that a complex is forming between GAPDH and PGK. This result in itself was encouraging, although the exact stoichiometry of the interaction could not be determined from this data. However, the scattering data from this experiment can be analyzed from a theoretical standpoint and a speculation as to the stoichiometry of the complex that is forming can be made. The following analysis is purely theoretical, although it is based on the scattering data obtained from the titration experiments.

It is reasonable to assume that since GAPDH is a tetramer, the stoichiometry of the binding of PGK to GAPDH will vary between a ratio of 1:1 and 1:4, GAPDH to PGK. Based on the scattering from one PGK, one GAPDH and the titration data itself, it is

possible to extrapolate the scattering data to each of the stoichimetric scenarios, in an attempt to determine which ratio of PGK to GAPDH best fits the observed data. For each stoichiometry, the analysis of I(0)'s alone will generate the quantity of protein that has formed the conjectured complex. Since the concentrations of each protein alone are known, the association constant K_a can be calculated. This is expected to be a constant for all concentrations studied, and so a flat line is expected when K_a is plotted against [PGK]/[GAPDH], if the correct stoichiometry is met. Variations from a straight line, that is, variations in K_a at different ratios of PGK to GAPDH, indicate that the conjectured complex does not adequately reflect the scattering data. If a truly flat line were observed, an R_G analysis would have been possible and would have yielded the distance from the center of mass of PGK to the center of mass of GAPDH. In turn, theoretical full scattering curves could then be generated and compared to the actual experimental data.

In order to determine truly the stoichiometry of binding, the NIST experiments need to be repeated. However, it is possible to speculate on the results of these studies:

There are three possible outcomes from these experiments. In Case 1, a 1:1 dPGK:GAPDH association is occurring. In this case, there would be no observed change in the R_G or scattering intensity of the system. This is because the GAPDH is matched out and effectively silent. In a case such as this, the scattering from one bound PGK would appear the same as if the PGK were free in solution. It is not expected that the PGK would adopt a closed conformation, although any conformational change would be be manifested as a change in the R_G. The R_G of PGK in the closed conformation is

known (Henderson *et al.*, 1994), and if such a change occurred, it could be analyzed in the context of this previous work.

In Case 2, a 2:1 dPGK:GAPDH association is occuring. In this case, some of the dPGK is free and some is bound to GAPDH. The scattering intensity of bound dPGK would be twice as much as the unbound form and the R_G of the bound dPGK would be approximately 55.5. This value assumes a 50 Å distance between the center of mass of the bound dPGK to the center of mass of the GAPDH, and the R_G of PGK to be 24.0 Å which was determined previously. The R_G of the complex can be described by

$$R_G^2{}_{complex} = R_G^2{}_{PGK\ monomer} + d^2 \quad \text{Eqn V-4}$$

where d is the distance between the centers of mass of the two proteins.
The distance of 50 Å is a rough estimate of distance based on docking studies of the crystal structures of yeast PGK and lobster GAPDH. If this situation were observed, it would be clearly distinguishable.

In Case 3, a 4:1 dPGK:GAPDH association is occuring. Assuming the dPGKs bind symetrically, the R_G of the complex would be the same as that determined for the 2:1 complex. However this case is distinguishable from Case 2 because the observed scattering intensity will be four times greater.

Preliminary results from this study suggest that there is indeed a complex forming between the two enzymes, although the stoichiometry of the association is unclear. The symmetry of GAPDH would seem to suggest a stoichiometry of either 2:1 or 4:1, PGK to GAPDH, for the complex that is forming. In order to confirm this absolutely, however,

the NIST contrast matching studies need to be attempted again, using freshly prepared deuterated PGK. Hopefully, the results of these future studies will not only provide further insight into this problem, but also demonstrate the utility of the deuterated enzyme in attempting similarly challenging work.

LIST OF REFERENCES

List of References

Adams, B., Burgess, R. J. & Pain, R. H. (1985) *Eur. J. Biochem.* **152**, 715-720.

Ashmarina, L. I., Muronetz, V. I. & Nagradova, N. K. (1984) *Biochem. Int.* **9**, 511-521.

Banks, R. D., Blake, C. C. F., Evans, P.R., Haser, R., Rice, D. W., Hardy, G. W., Merrett, M. & Phillips, A .W. (1979) *Nature* **279**, 773-777.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Magn. Res.* **55**, 301.

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

Bendall, M. R., Doddrell, D. M. & Pegg, D. T. (1981) *J. Am. Chem. Soc.* **103**, 4603-4605.

Bennet, W. S. & Steitz, T. A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4848-4852.

Bernstein, B. E., Michels, P. A. M. & Hol, W. G. J. (1997) *Nature* **385**, 275-278.

Blake, C. C. F. & Evans, P. R. (1974) *J. Mol. Biol.* **84**, 585-601.

Blake, C. C. F. & Rice, D. W. (1981) *Phil. Trans. R. Soc. Lond. [Biol]* **293**, 93-104.

Bloch, W., MacQuarrie, R. A. & Bernhard, S. A. (1971) *J. Biol. Chem.* **246**, 780-790.

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

Bryant, T. N., Watson, H.C. & Wendell, H.C. (1974) *Nature* **247**, 14-17.

Bucher, T. (1942) *Biochim. Biophys. Acta* **1**, 292-314.

Chan, T. M. & Markley, J. L. (1982) *J. Am. Chem. Soc.* **104**, 4010-1018.

Clarke, F. M. & Masters, C. J. (1973) *Biochim. Biophys. Acta* **327**, 223-226.

Crespi, H. L., Conrad, S. M., Uphaus, R. A. & Katz, J. J. (1960) *Ann. N.Y. Acad. Sci.* **84**, 648-666.

Crespi, H. L., Rosenberg, R. M. & Katz, J. J. (1968) *Science* **161**, 795-796.

Crespi, H. L., Daboll, H. F. & Katz, J. J. (1970) *Biochim. Biophys. Acta* **200**, 26-33.

Crespi, H. L., Smith, L., Gajda, T. & Ammeraal, R. M. (1972) *Biochim. Biophys. Acta* **256**, 611-618.

Datta, A. Merz, J. M. & Spivey, H. O. (1985) *J. Biol. Chem.* **260**, 15008-15012.
Davies, G. J., Gamblin, S. J., Littlechild, J. A., Dauter, Z., Wilson, K. S. & Watson, H. C. (1994) *Acta Crystallog.* **D50**, 202-209.

DiGiammarino, E. L. (1996) *M.S. Thesis*, University of Tennessee, Knoxville.

Dryden, D. T. F., Varley, P. G. & Pain, R. G. (1992) *Eur. J. Biochem.* **208**, 115-123.

Fifis, T. & Scopes, R.K. (1978) *Biochem. J.* **175**, 311-319.

Friedrich, P. & Hajdu, J. (1987) *Biochem. Soc. Transact.* **15**, 973-977.

Gibson, A. (1997) unpublished protocol.

Gregory, J. D. (1993) *Ph.D. Dissertation*, University of Tennessee, Knoxville.

Gregory, J. D., & Serpersu, E. H. (1993) *J. Biol. Chem.* **268**, 3880-3888.

Harlos, K., Vas, M. & Blake, C. F. (1992) *Proteins Struc. Funct. Genet.* **12**, 133-144.

Harpel, M. R., Serpersu, E. H., Lamerdin, J. A., Huang, Z.-H., Gage, D. A. & Hartman (1995) *Biochemistry* **34**, 11296-11306.

Henderson, S. J., Serpersu, E. H., Gerhardt, B. S. & Bunick, G. J. (1994) *Biophys. Chem.* **53**, 95-104.

Horecker, B. L., Melloni, E. & Pontremoli, S. (1981) *Curr. Top. Cell. Regul.* **18**, 181-197.

Ikura, M., Marion, D., Kay, L. E., Shih, H., Krinks, M., Klee, C. B. & Bax, A. (1990a) *Biochem. Pharmacol.* **40**, 153-160.

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

Jeener, J. (1971) *Ampere Inernational Summer School*, Basko Polje, (proposal).

Kalbitzer, H. R., Leberman, R. & Wittinghofer, A. (1985) *FEBS Lett.* **180**, 40-42.

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

Keleti, T. & Ovadi, J. (1988) *Curr. Top. Cell. Regul.* **29**, 1-33.

Keleti, T., Batke, J., Ovadi, J., Jancsik, V. & Bartha, F. (1977) *Adv. Enzyme Regul.* **15**, 233-265.

Kenney, F. T. (1959) *J. Biol. Chem.* **234**, 2707-2712.

Khoroshilova, N. A., Muronetz, V. I. & Nagradova, N. K. (1992) *FEBS Lett.* **297**, 247-249.

Kuntz, G. W. K. & Kreitsch, W. K. G. (1982) *Methods Enzymol.* **90**, 110-114.

Kvassman, J. & Pettersson, G. (1989) *Eur. J. Biochem.* **186**, 265-272.

LeMaster, D. M. (1990) *Quarterly Rev. Biophys.* **23**, 133-174.

LeMaster, D. M. (1989) *Methods Enzymol.* **177**, 23-43.

LeMaster, D. M. & Richards, F. M. (1988) *Biochemistry* **27**, 142-150.

Lerner, L. & Bax, A. (1986) *J. Magn. Res.* **69**, 375-380.

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

Markland, F. S., Bacharach, A. D. E., Weber, B. H., O'Grady, T. C., Saunders, G. C. & Umemura, N. (1975) *J. Biol. Chem.* **250**, 1301-1310.

Markley, J. L., Putter, I. & Jardetsky, O. (1968) *Science* **161**, 1249-1251.

Mastro, A. M. & Keith, A. D. (1984) *J. Cell Biol.* **99**, 1805.

McPhillips, T. M., Hsu, B. T., Sherman, M. A., Mas, M. T. & Rees, D. C. (1996) *Biochemistry* **18**, 338-342.

Mohan, V. S., Crespi, H. L. & Katz, J. J. (1962) *Nature* **193**, 189-190.

Morris, G. A. & Freeman, R. (1979) *J. Am. Chem. Soc.* **101**, 760-762.

Pappu, K. M., Kunnumal, B. & Serpersu, E. H. (1997) *Biophys. J.* **72**, 928-935.

Pappu, K. M., Gregory, J. D. & Serpersu, E. H. (1994) *Arch. Biochem. Biophys.* **311**, 503.

Pappu, K. M. (1996) *Ph.D. Dissertation*, University of Tennessee, Knoxville.

Pappu, K. M. & Serpersu, E. H. (1994) *J. Magn. Res. Series B* **105**, 157-166.

Pickover, C. A., McKay, D. B., Engleman, D. B. & Steitz, T. A. (1979) *J. Biol. Chem.* **254**, 11323-11329.

Roustan, C., Fattoum, A., Jeanneau, R. & Pradell, L. A. (1980) *Biochemistry* **34**, 833-841.

Roustan, C., Fattoum, A. & Pradel, L.-A. (1976) *Biochemistry* **15**, 2172-2177.

Roustan, C., Fattoum, A. & Pradel, L.-A. (1979) *Biochimie* **61**, 663-669.

Sanders, J. K. M. & Hunter, B.K. (1994) Modern NMR Spectroscopy: A Guide For Chemists. Second Edition. Oxford University Press Inc, New York.

Scopes, R. K. (1971) *Biochem. J.* **122**, 89-92.

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

Shibata, C. G., Gregory, J. D., Gerhardt, B. S. & Serpersu, E. H. (1995) *Arch. Biochem. Biophys.* **319**, 204-210.

Srere, P. A. (1981) *Trends Biochem. Sci.* **6**, 4.

Srivastava, D. K. & Bernhard, S. A. (1987) *Ann. Rev. Biophys. Biophys. Chem.* **16**, 175-204.

Srivastava, D. K. & Bernhard, S. A. (1985) *Biochemistry* **24**, 623-628.

Sukhodolets, M. V., Muronetz, V. I., Tsuprun, V. L., Kaftanova, A. S. & Nagradova, N. K. (1988) *FEBS Lett.* **238**, 161-166.

Sukhodolets, M. V., Muronetz, V. I., Nagradova, N. K. (1989) Biochem. *Biophys. Res. Comm.* **161**, 187-196.

Tanswell, P., Westhead, E. W. & Williams, R. J. P. (1976) *Eur. J. Biochem.* **63**, 249-262.

Vas, M. & Batke, J. (1990) *Eur. J. Biochem.* **191**, 679-683.

Vas, M. & Batke, J. (1981) *Biochim. Biophys. Acta* **660**, 193-198.

Volkman, B. (1996) NMRFAM protocols.

Wagner, G. & Wuthrich, K. (1982) *J. Mol. Biol.* **155**, 347-366.

Warwicker, J. (1989) *J. Mol. Biol.* **206**, 381-395.

Watson, H. C., Walker, N. P. C., Shaw, P. J., Bryant, T. N., Wendell, P. L., Fothergill, L. A., Perkins, R. E., Conroy, S. C., Dobson, M. J., Tuite, M. F., Kingsman, A. J. & Kingsman, S. M. (1982) *EMBO J.* **1**, 1635-1640.

Webb, M. R. & Trentham, B. R. (1980) *J. Biol. Chem.* **255**, 1775-1779.

Weber, J. P. & Bernhard, S. A. (1982) *Biochemistry* **21**, 4189-4194.

Wishart, D. S., Sykes, B. D. & Richards, F. M. (1993) *Biochim. Biophys. Acta* **1164**, 36-46.

Yamazaki, T., Tochio, H., Furui, J., Aimoto, S. & Kyogoku, Y. (1997) *J. Am. Chem. Soc.* **119**, 872-879.

Appendices

Appendix A

Figures

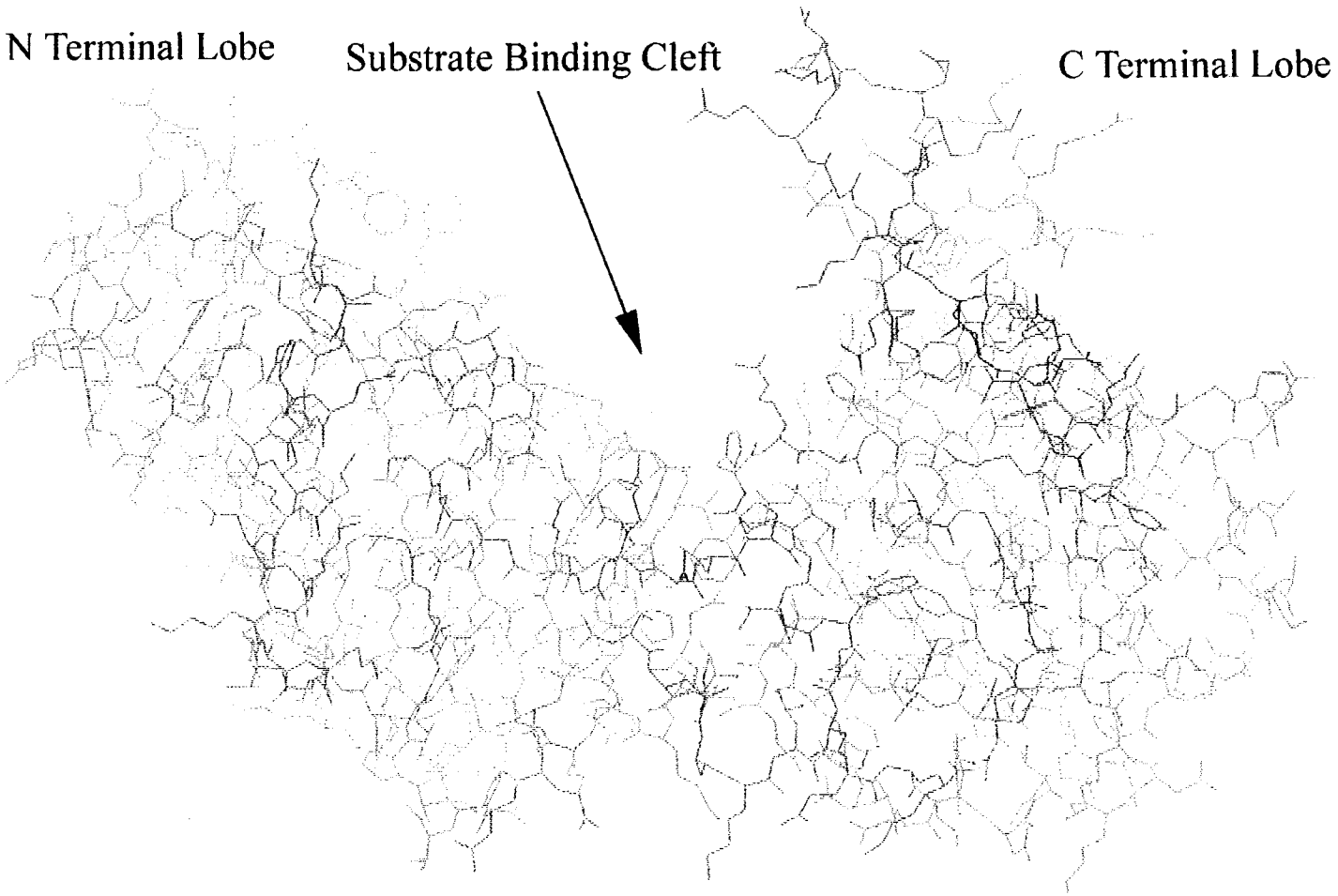


Figure I-1. Crystal strucutre of yeast phosphoglycerate kinase. The two lobes and the hinge region of the enzyme are clearly visible. The substrate binding cleft is indicated. (Watson *et al.*, 1982)

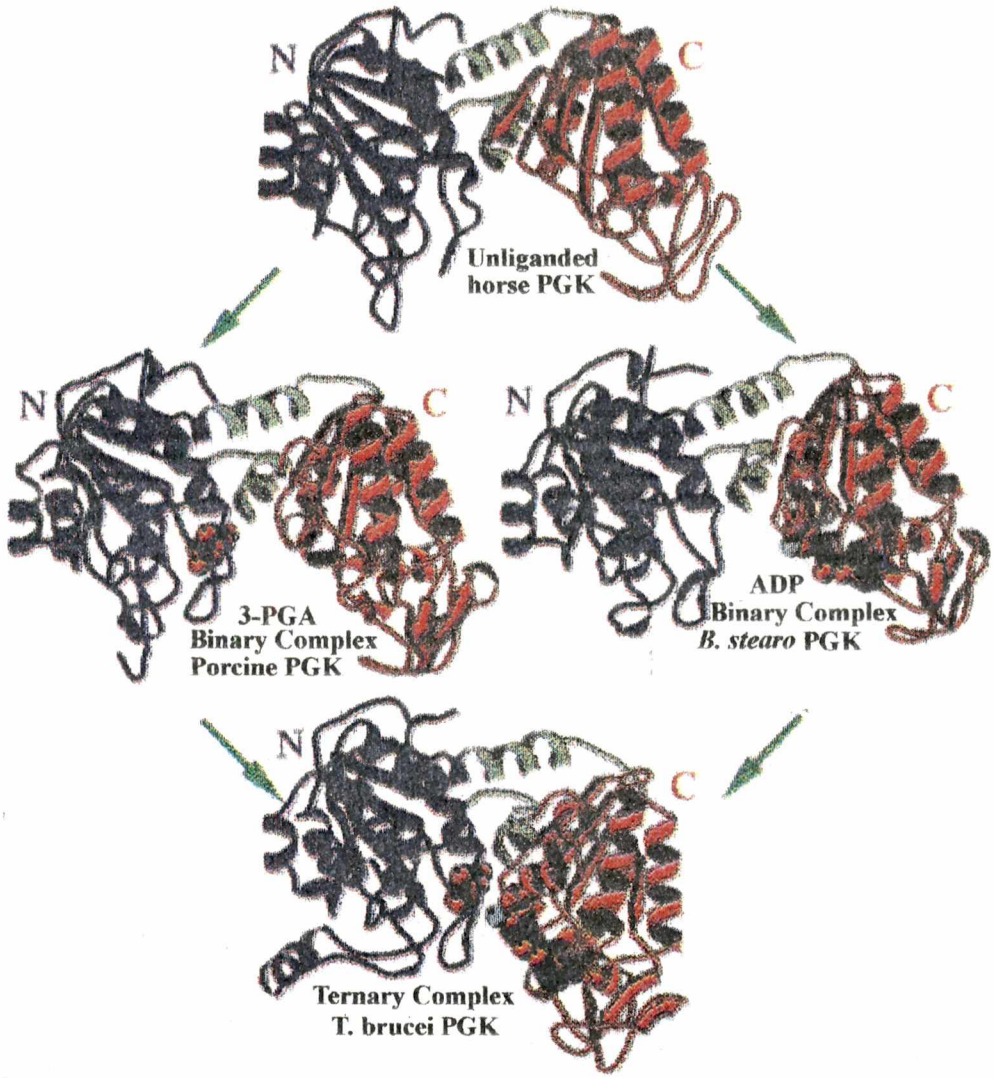


Figure I-2. The four different states of substrate-induced conformational change in PGK. The results of four different structural studies are collected in this model, which is the first depiction that includes the ternary complex. PGK binds its substrates in random order and therefore both binary complexes are shown as steps to the ternary complex. (Bernstein *et al.*, 1997)

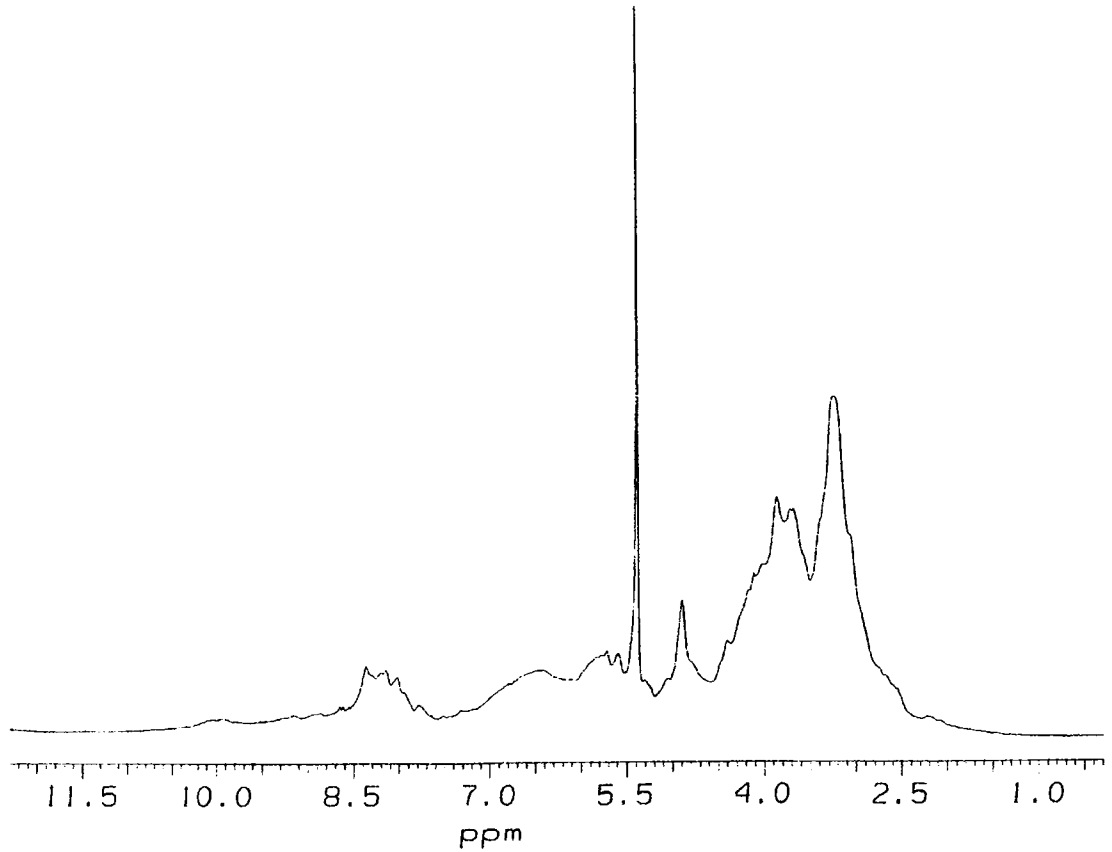


Figure II-1. Proton spectrum of isotopically normal PGK. Sample concentraton is approximately 0.5 mM.

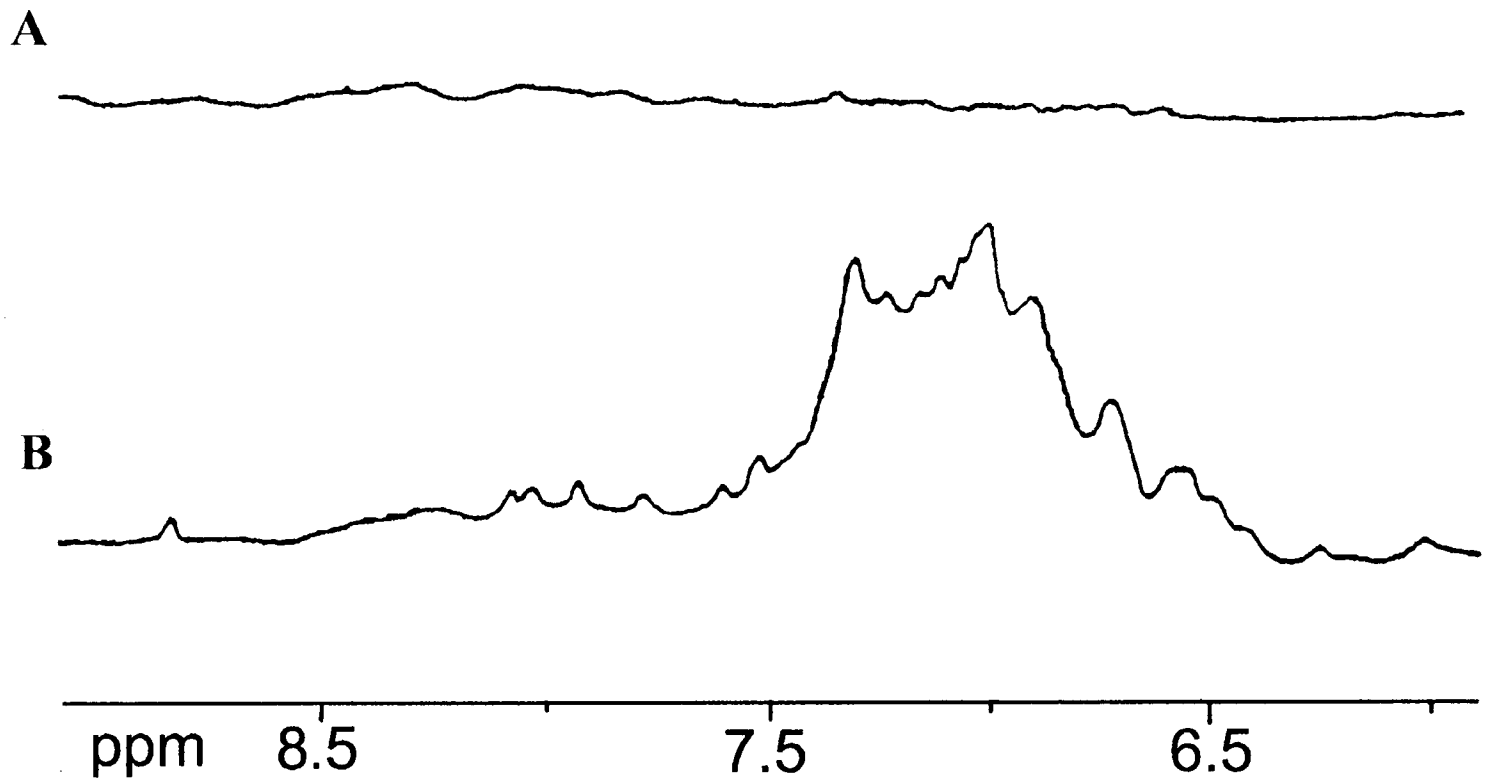


Figure II-2. Proton spectrum of the aromatic regions of perdeuterated yeast PGK (A) and isotopically normal yeast PGK (B). Sample concentraion of the protonated sample is approximately 0.5 mM. The concentration of the deuterated sample is approximately 0.3 mM.

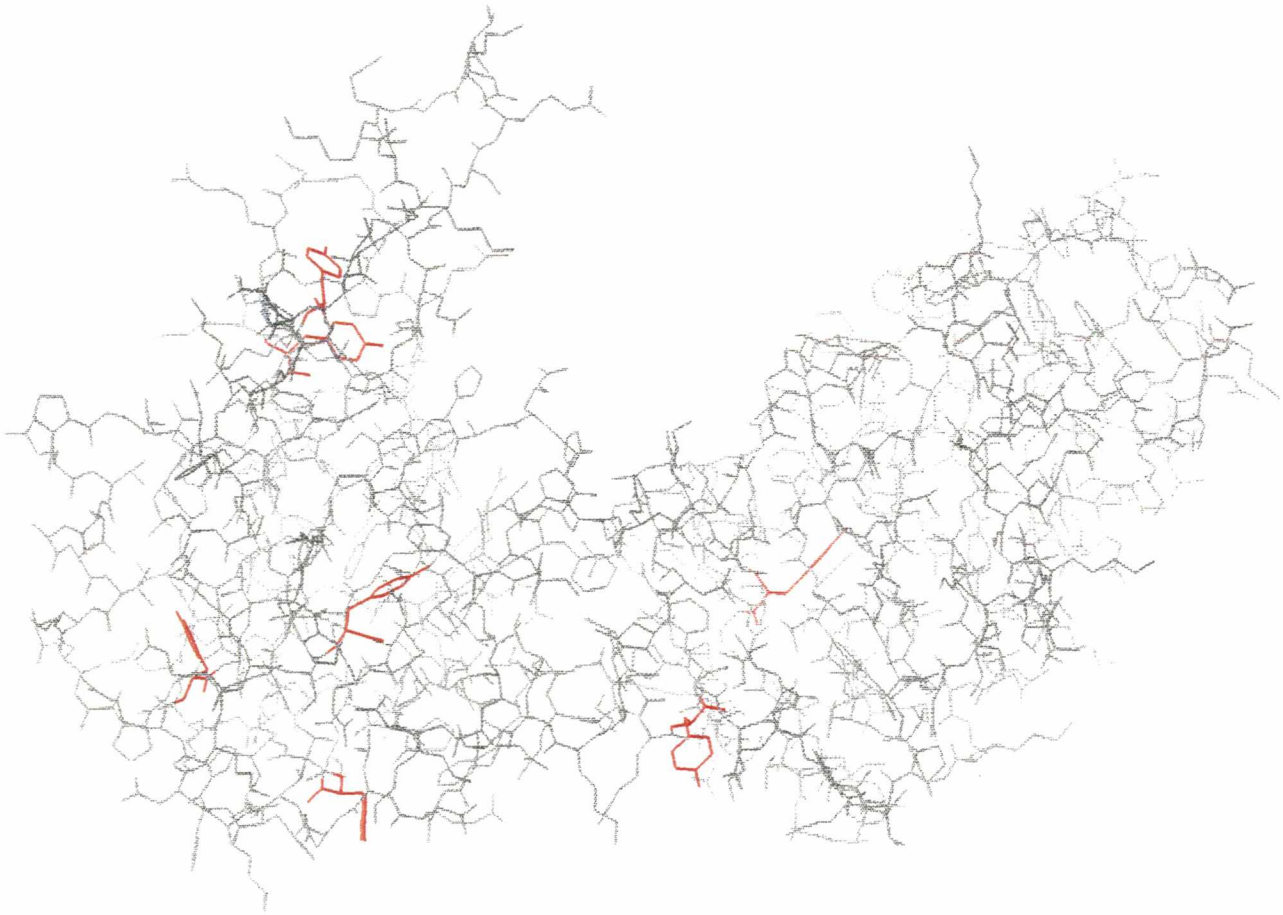


Figure II-3. Structure of yeast PGK showing the location of the seven tyrosine residues. found in the sequence of PGK. The tyrosine residues are highlighted in red. (Watson *et al.*, 1982)

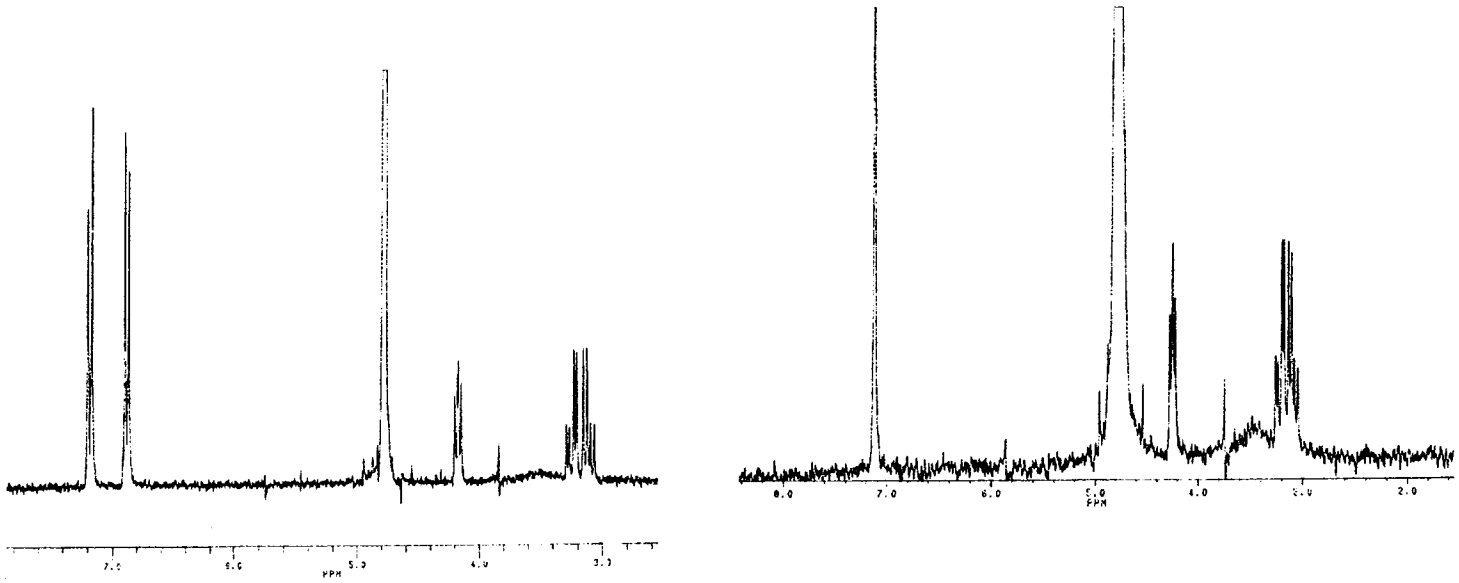


Figure II-4. Proton spectra of protonated (left) and selectively deuterated (right) tyrosine. The concentration of both samples is approximately 10 mM. Both spectra were processed identically, with the exception that the spectrum of the protonated sample was not baseline corrected.

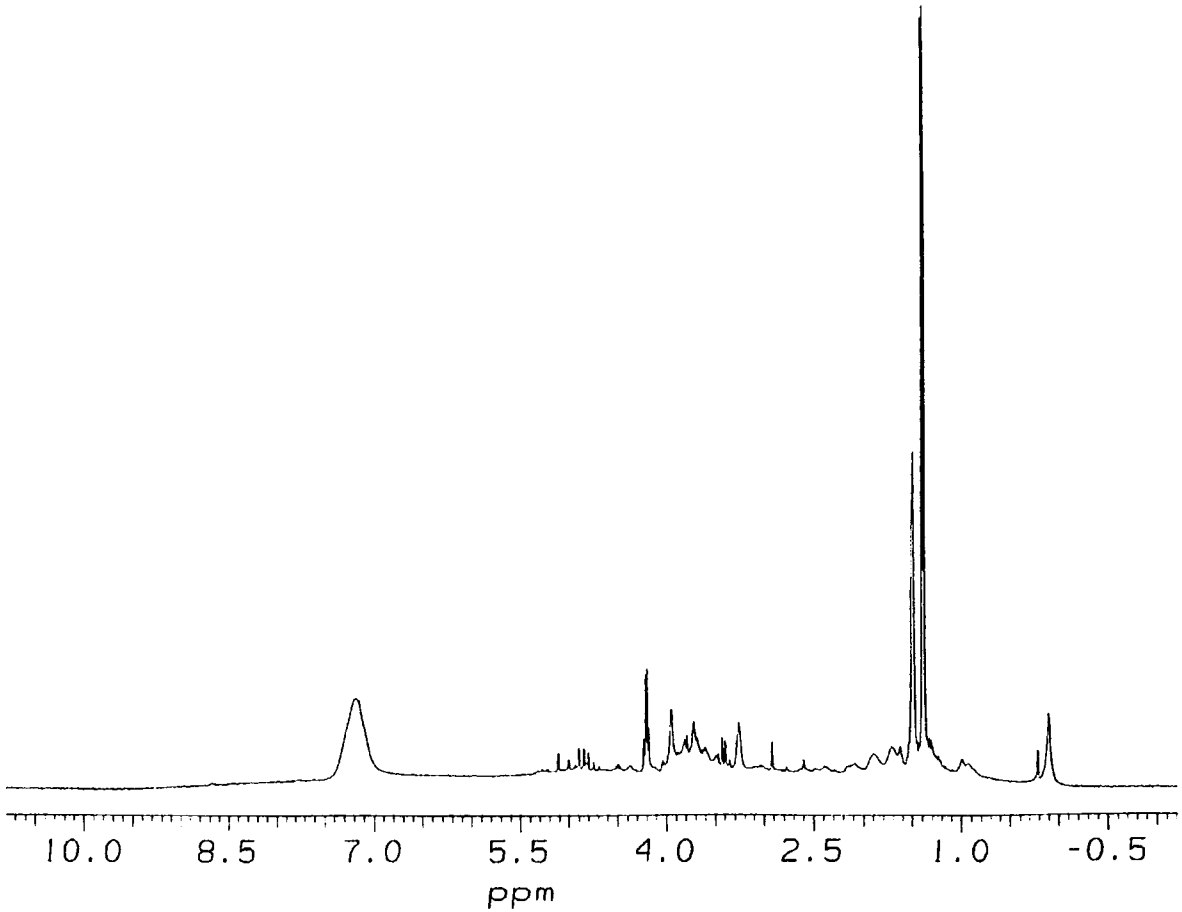


Figure II-5. Proton spectrum of selectively deuterated tyrosine-labeled dPGK. The concentration of the sample was approximately 0.4 mM.

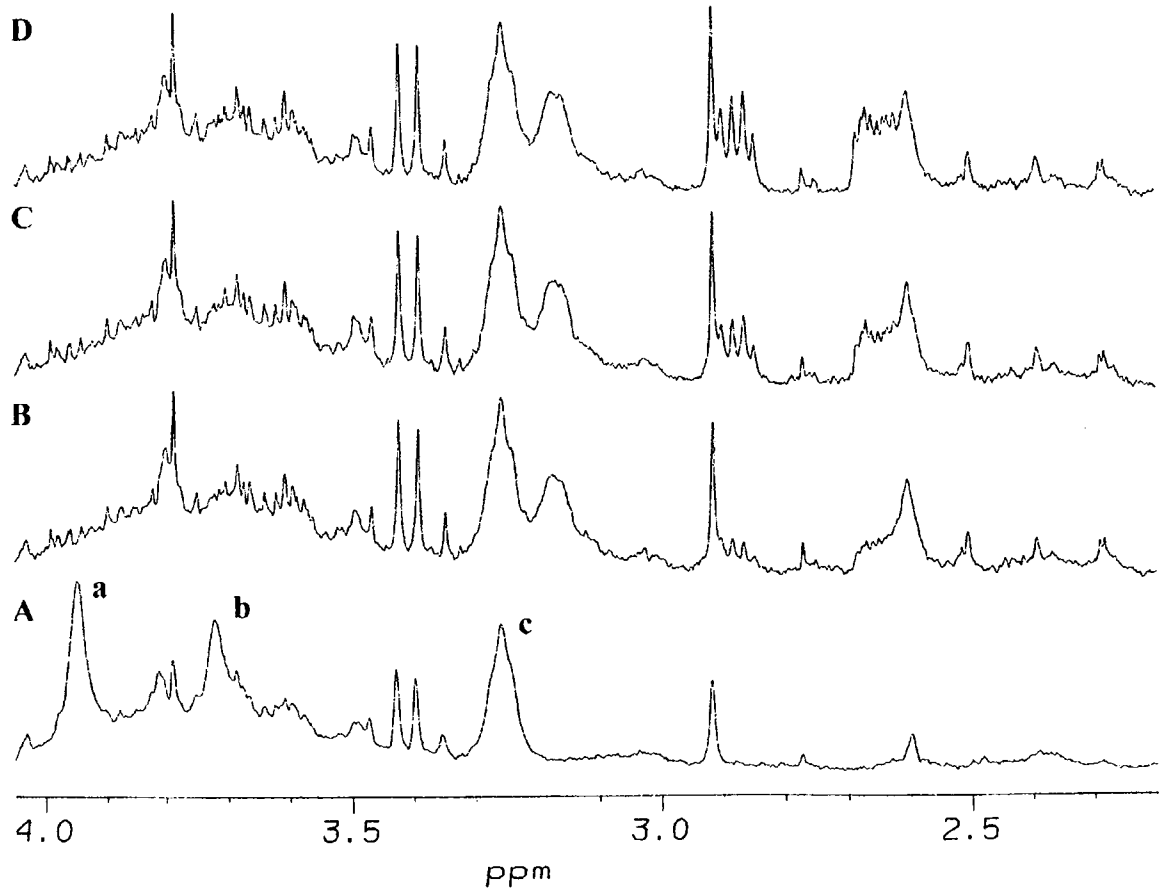


Figure II-6. Effect of the addition of MgATP on the proton spectrum of ^{1}H Tyr dPGK. A: No MgATP added; B: MgATP:enzyme ratio 0.5:1; C: MgATP:enzyme ratio 1:1; D: MgATP:enzyme ratio 1.5:1. Enzyme concentration was approximately 0.4 mM. Peaks effected by the addition of substrate are labeled.

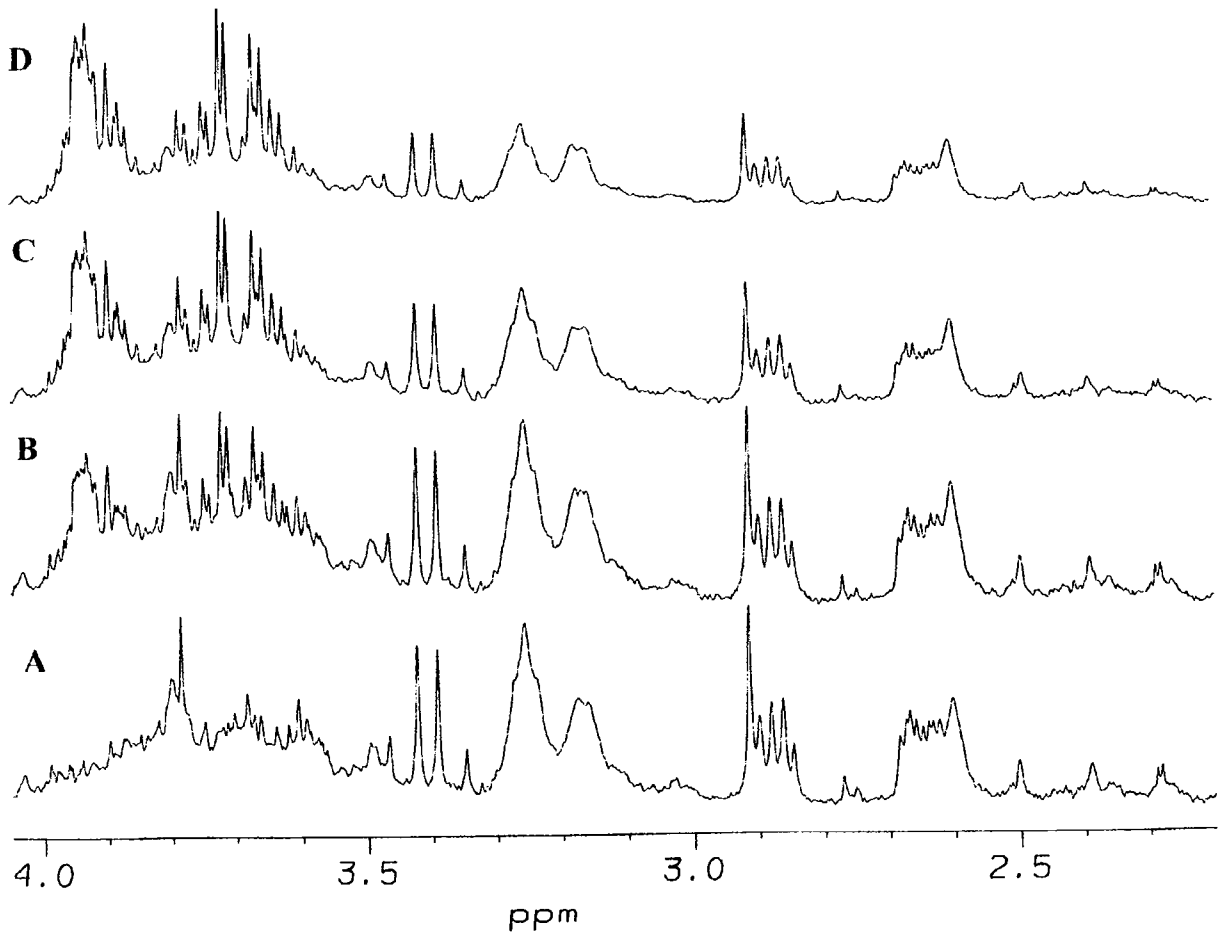


Figure II-7. Effect of the addition of glycerol-3-phosphate (G3P) on the binary MgATP/^{1}H Tyr dPGK complex. A: MgATP/enzyme, no G3P added, B: G3P:MgATP/dPGK ratio 0.5:1,C:G3P:MgATP/enzyme ratio 1:1, D: G3P:MgATP/enzyme ratio 1.5:1. Enzyme concentration was approximately 0.4 mM.

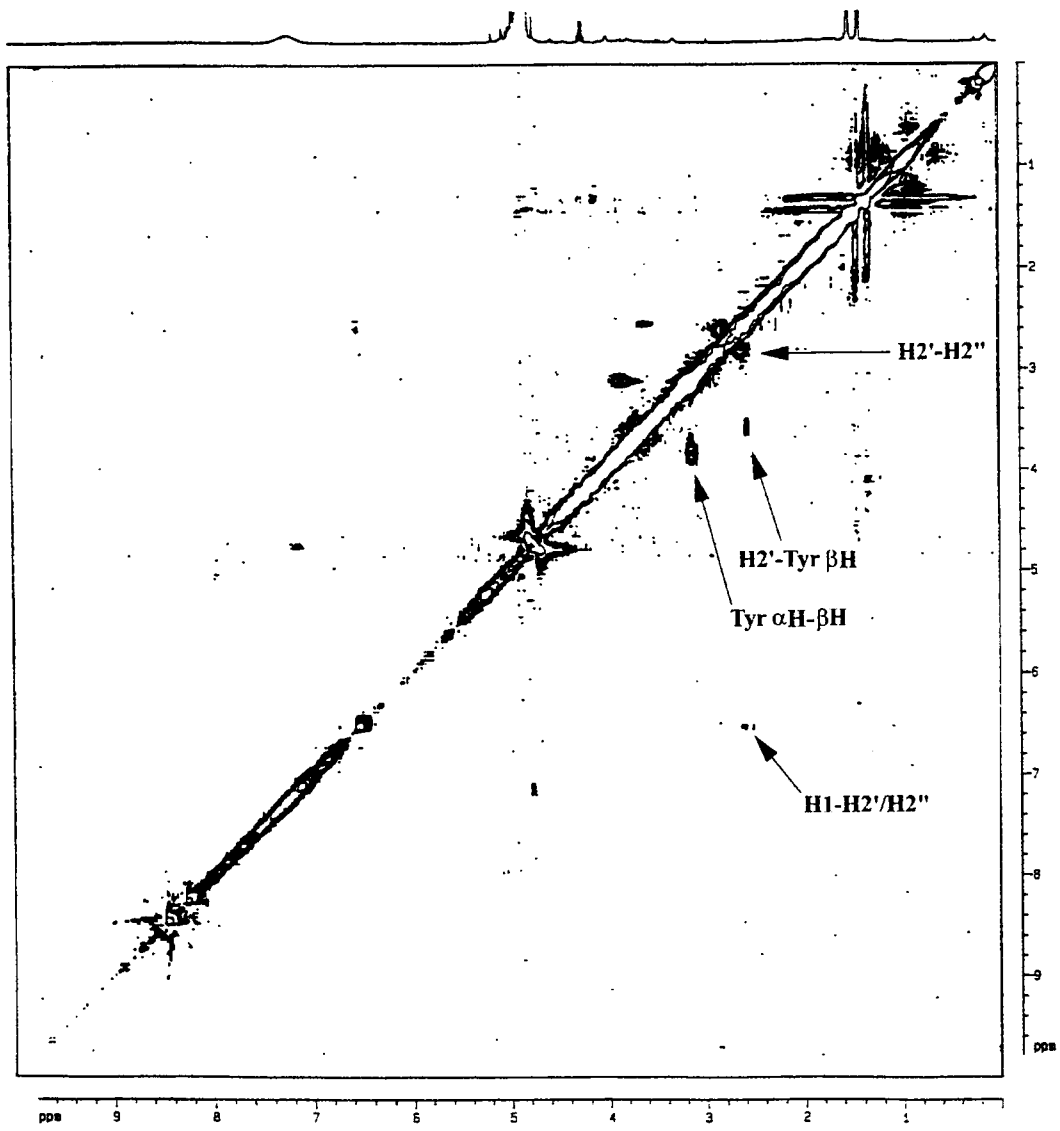


Figure II-8. NOESY spectrum of the ternary complex. The ratio of MgATP and G3P to ^{1}H Tyr dPGK is 1.5:1. The intermolecular NOE from the β proton of tyrosine to the H2' proton of ATP is indicated, as are other intramolecular NOEs. Enzyme concentration was approximately 0.4 mM and the mixing time of the experiment was 200 msec. The spectrum was not symmetrized.

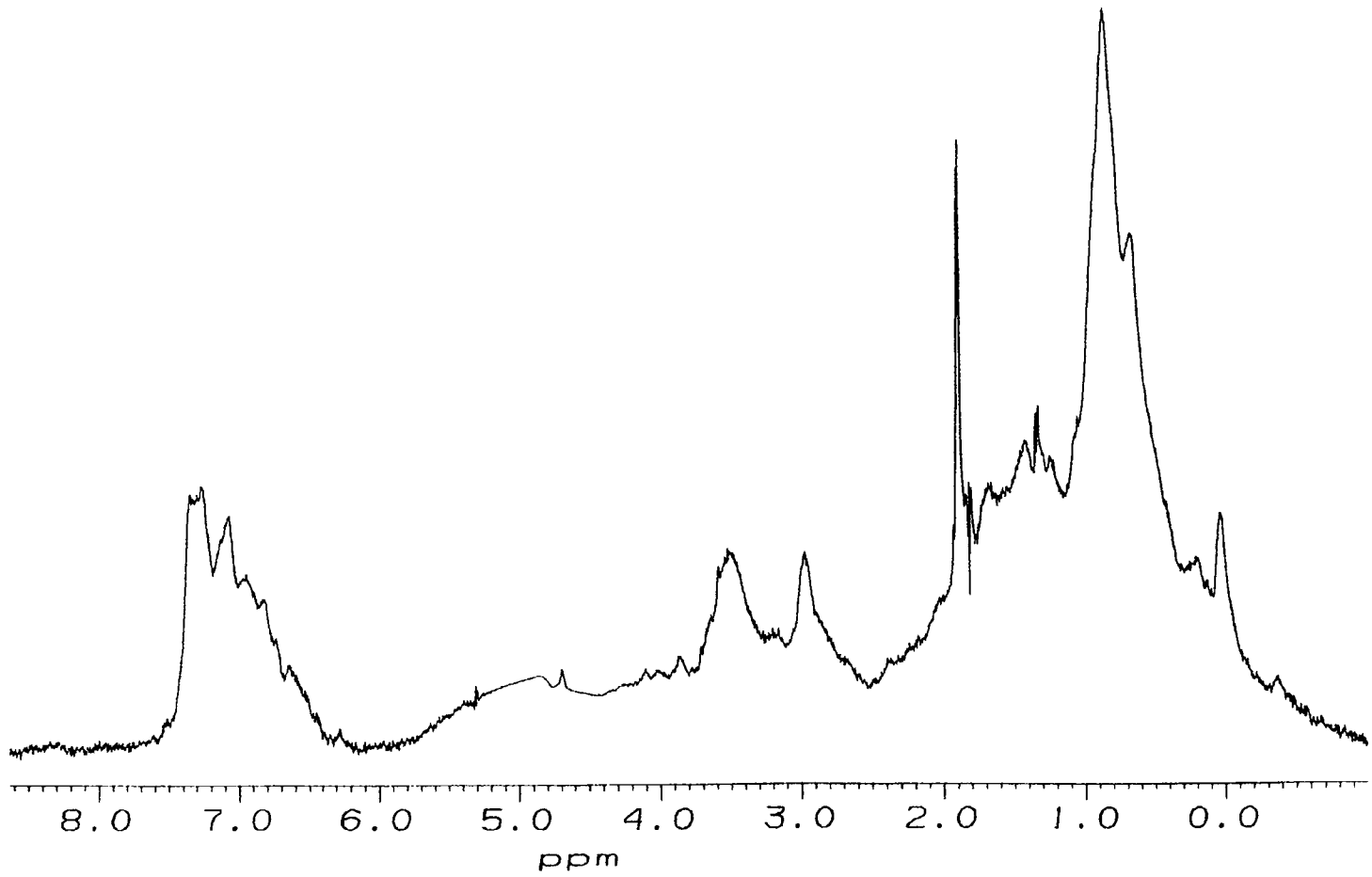


Figure III-1. Proton Spectrum of ^{15}N Phe ^{15}N Ile ^{13}C Tyr dPGK. Sample concentration was approximately 0.5 mM.

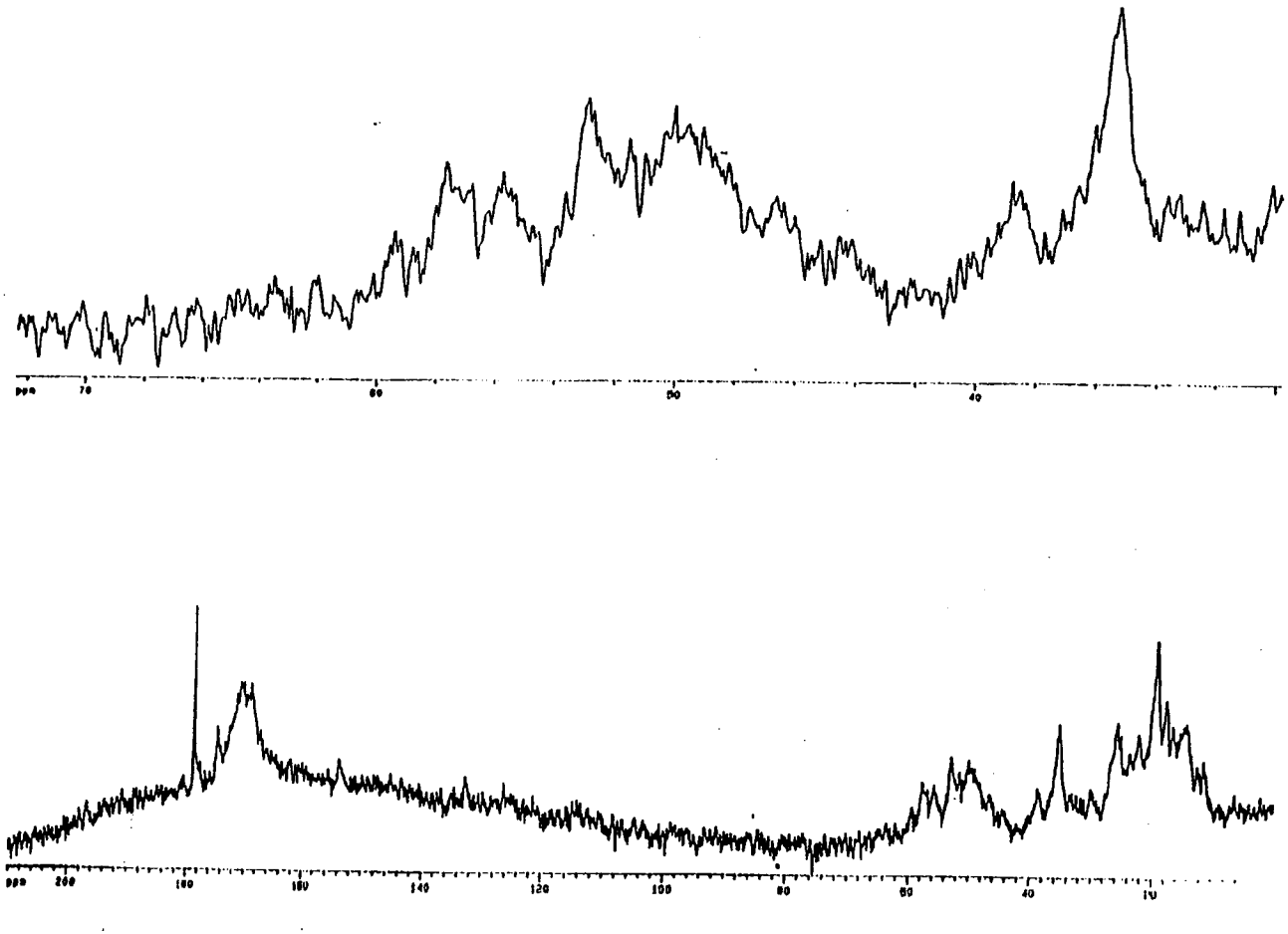


Figure III-2. ^{13}C spectrum of ^{15}N Phe ^{15}N Ile ^{13}C Tyr dPGK. Enzyme concentration was 0.5 mM. The lower trace covers the entire spectrum and the upper trace is an expansion of the alpha carbon region.

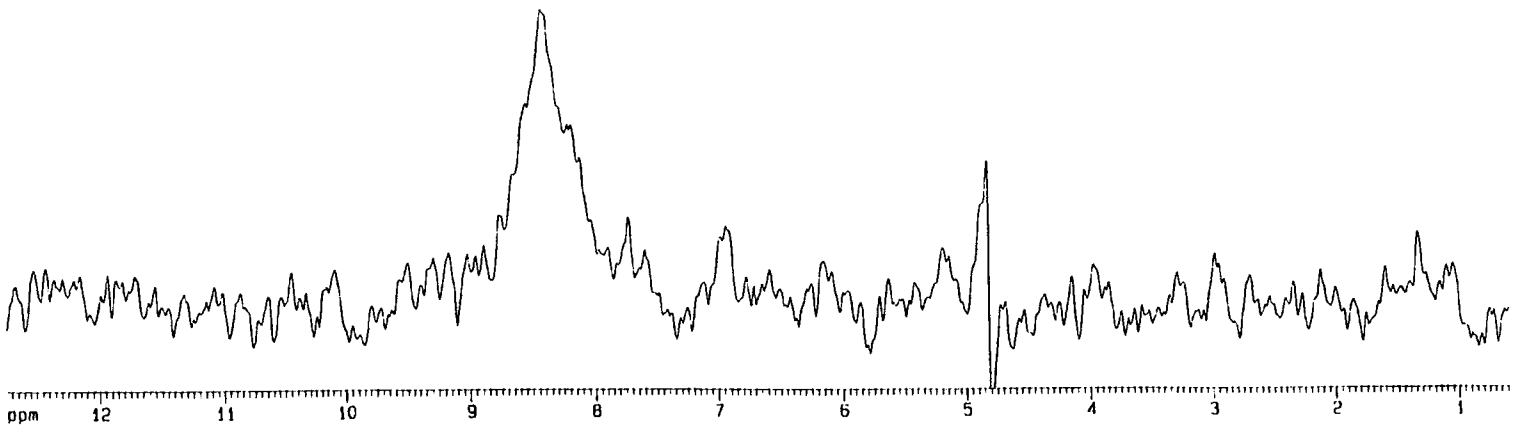


Figure III-3. X-filtered proton spectrum of ^{15}N Phe ^{15}N Ile ^{13}C Tyr dPGK. The enzyme concentration is 0.5 mM. This group of peaks around 8.5 ppm corresponds to the ^{15}N-bound protons in the enzyme.

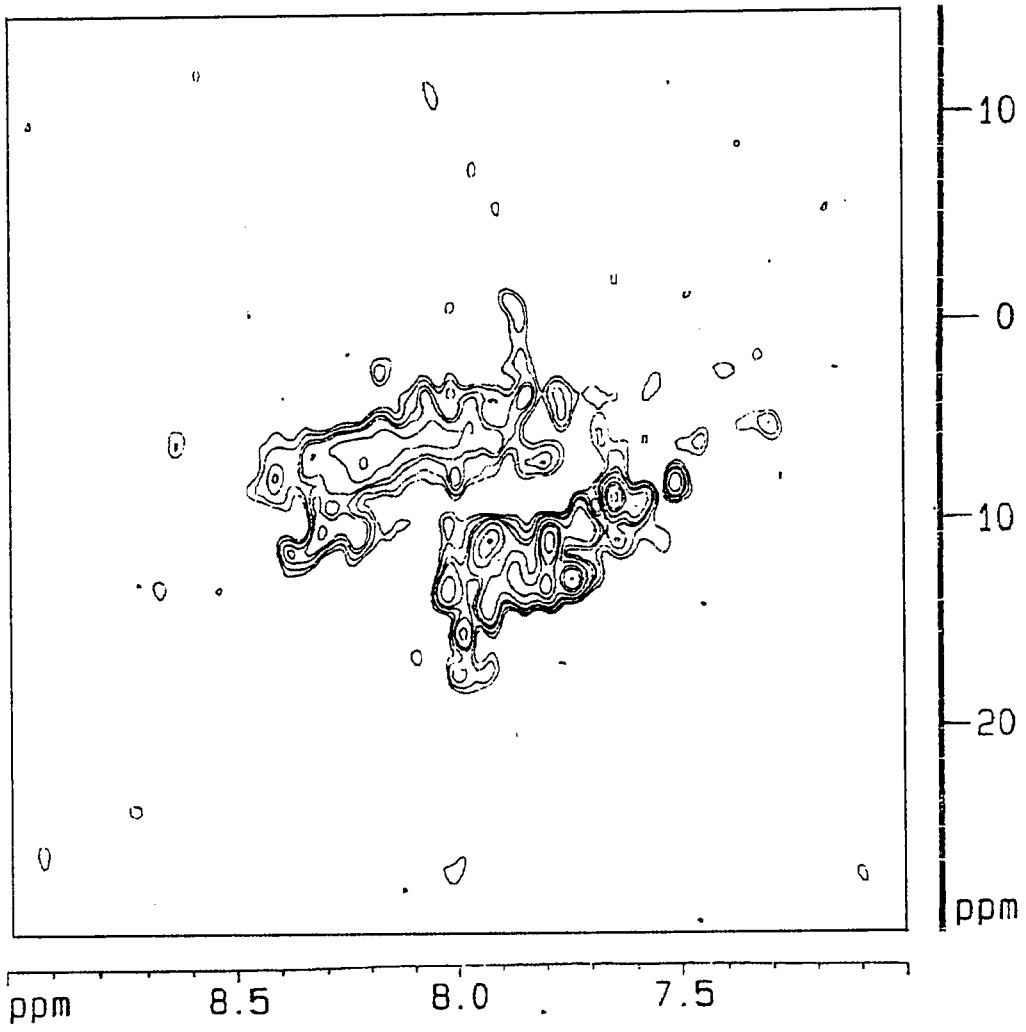


Figure III-4. Amine region of an HSQC spectrum of ^{15}N Phe ^{15}N Ile $^{13}C_{\alpha}$ Tyr dPGK. This spectrum correlates the ^{15}N label of phenylalanine and isoleucine and the protons bound to the labeled nitrogen. The vertical axis corresponds to the ^{15}N chemical shifts and the horizontal axis corresponds to the proton chemical shifts.

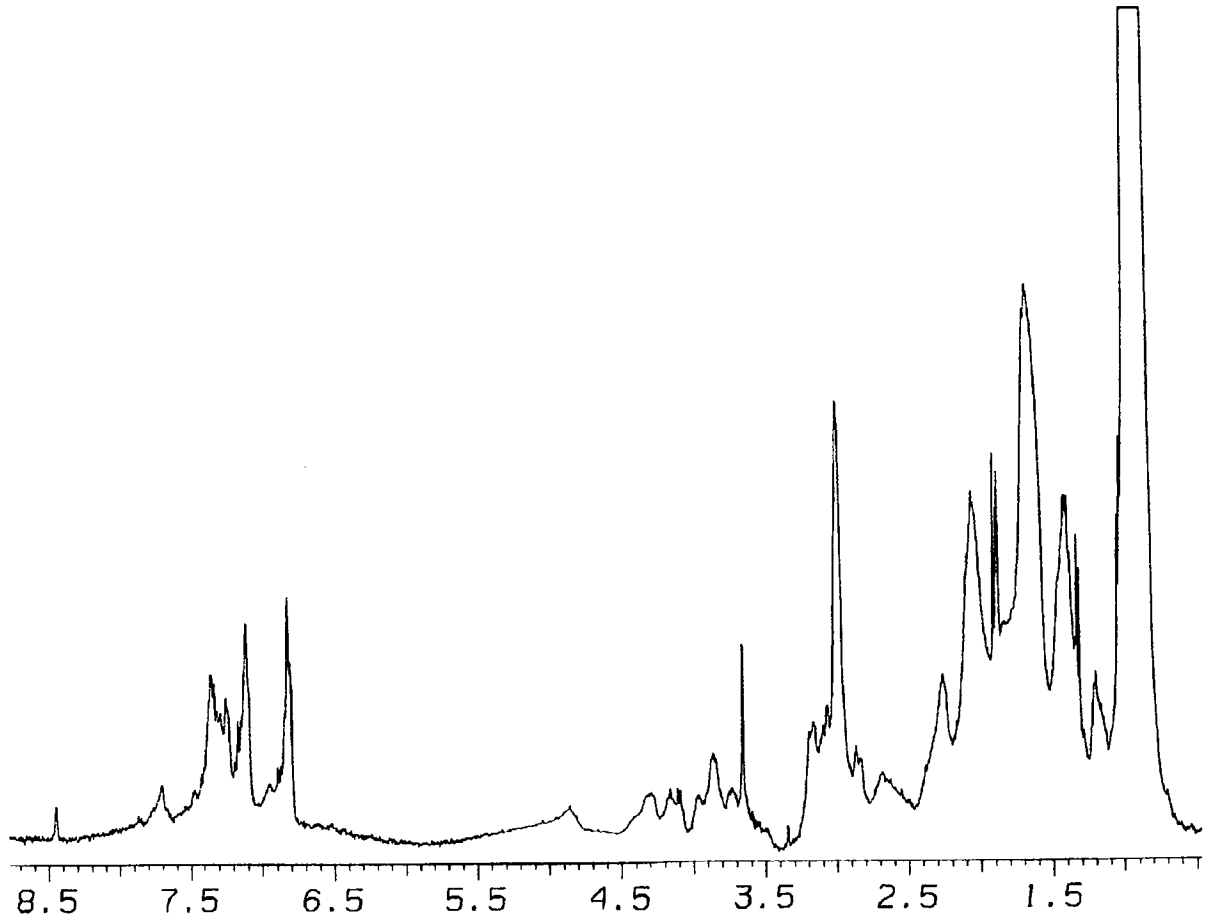


Figure III-5. Proton spectrum of ^{15}N Val ^{13}C Tyr dPGK. Protein concentration was 0.5 mM.

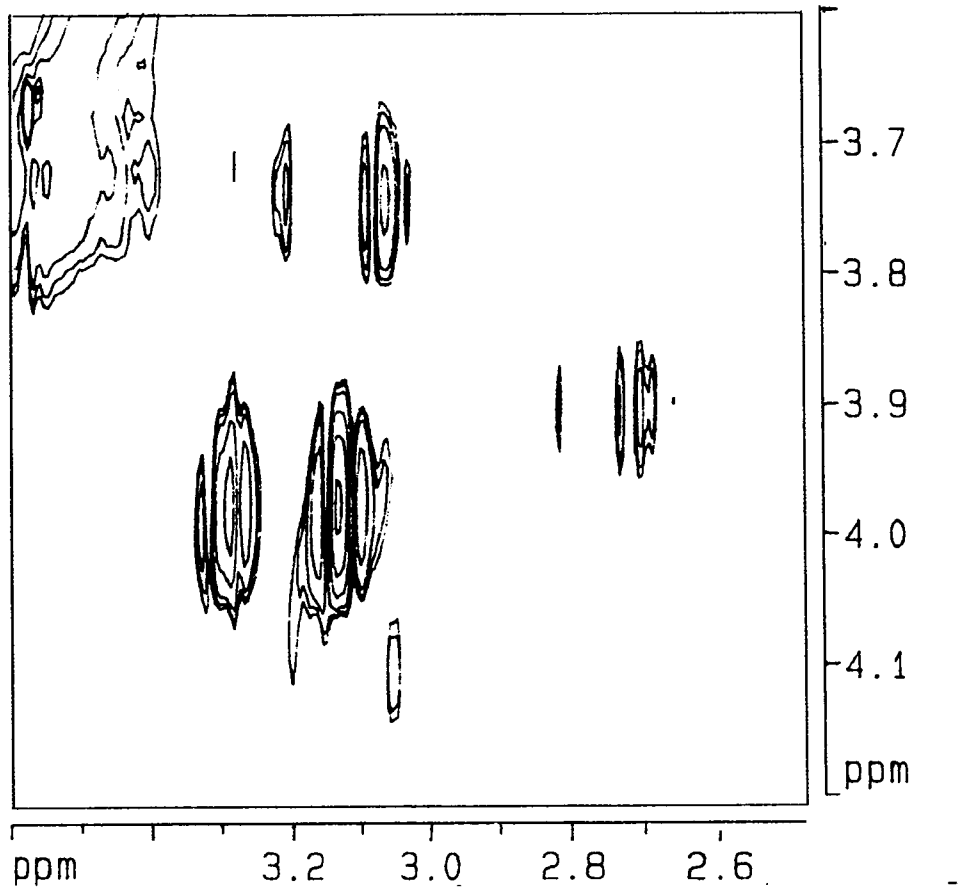


Figure III-6. TOCSY spectrum of ^{15}N Val ^{13}C Tyr dPGK. This region of the spectrum shows the αH-to-βH connectivities of tyrosine. Protein concentration was 0.5 mM.

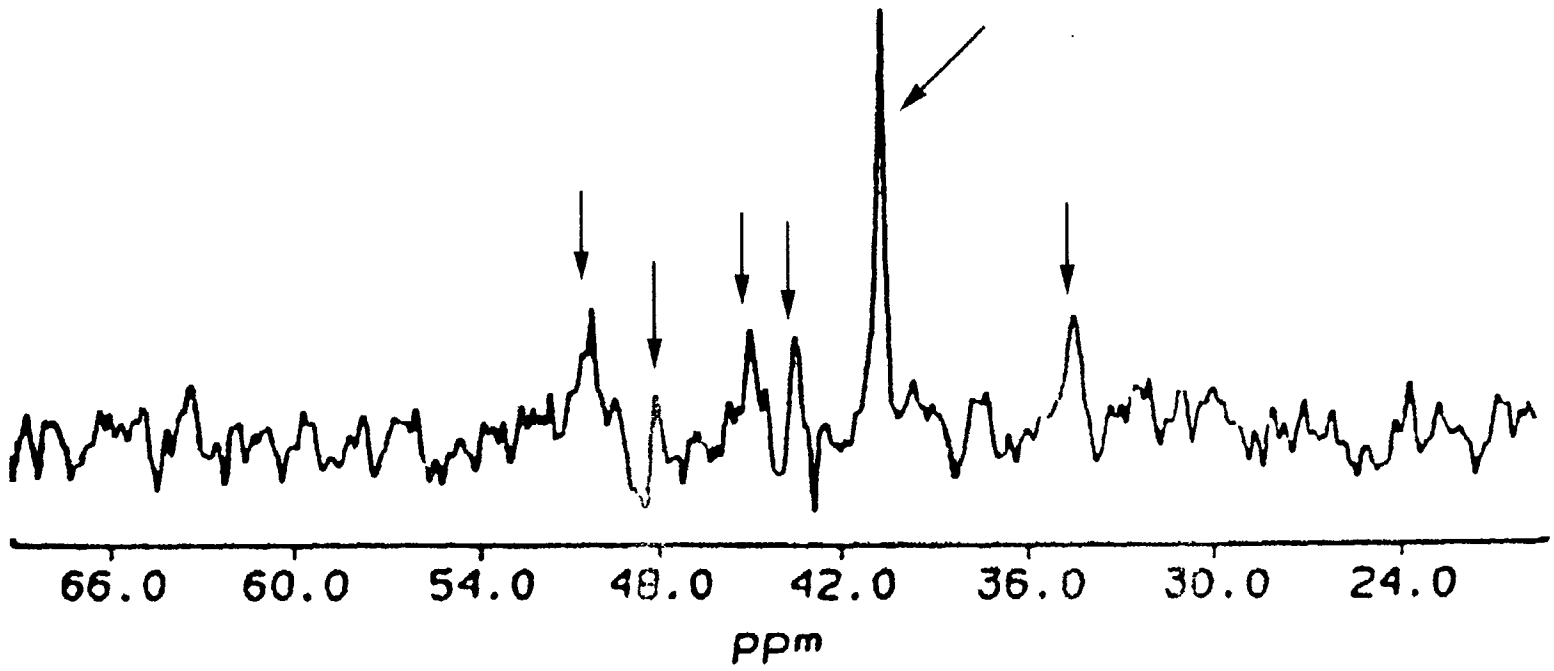


Figure III-7. ^{13}C DEPT spectrum of ^{15}N Val ^{13}C Tyr dPGK. The peaks corresponding to the $^{13}C_{\alpha}$ of tyrosine are indicated. The large peak is likely composed of two overlapping peaks. Protein concentraton was 0.5 mM.

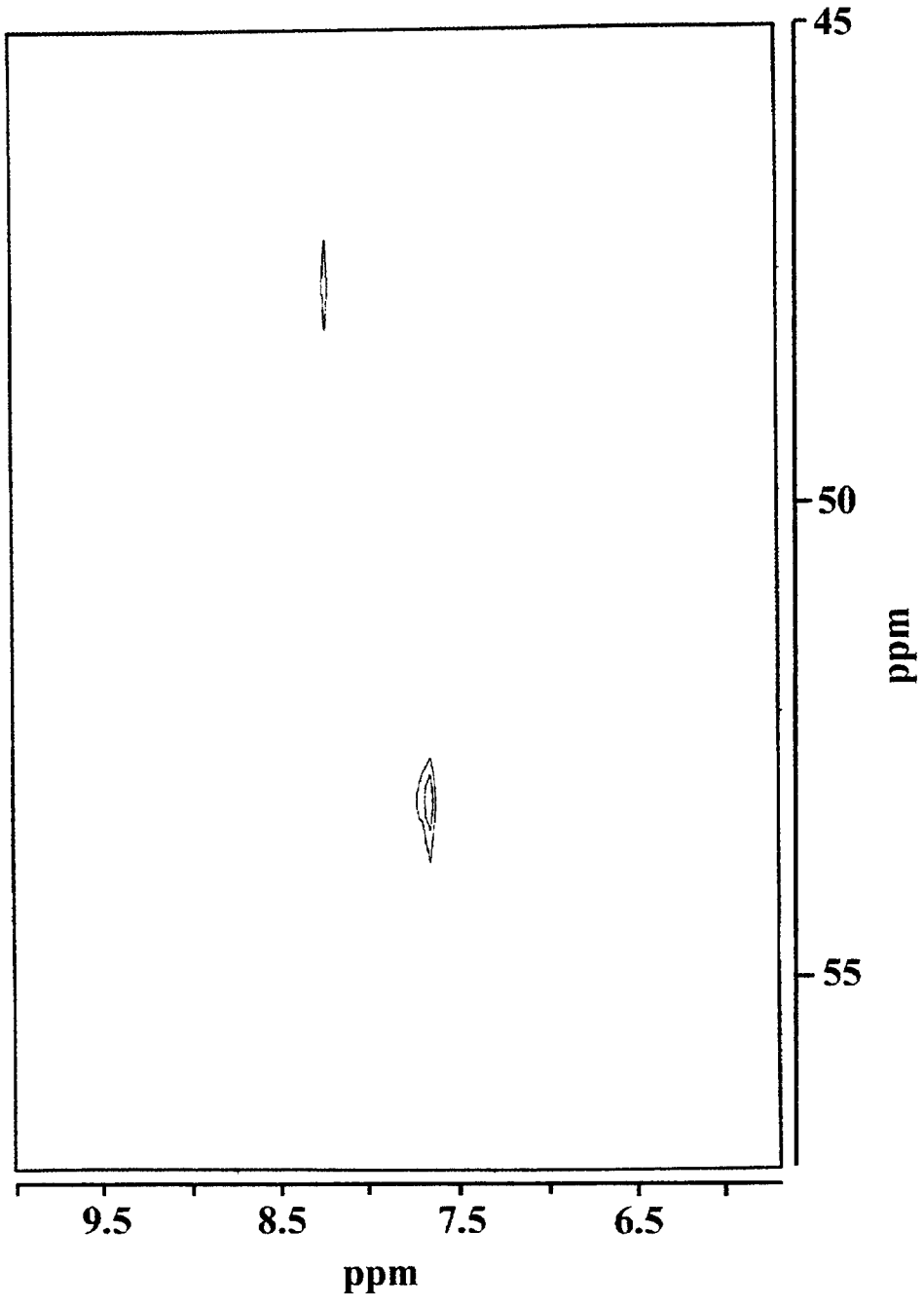


Figure III-8. HNCA spectrum of ^{15}N Val ^{13}C Tyr dPGK. The crosspeaks corresponding to interactions between the ^{15}N-bound proton of valine and the $^{13}C_{\alpha}$ of tyrosine are observed. The horizontal axis is the proton dimension and the vertical axis is the carbon dimension.

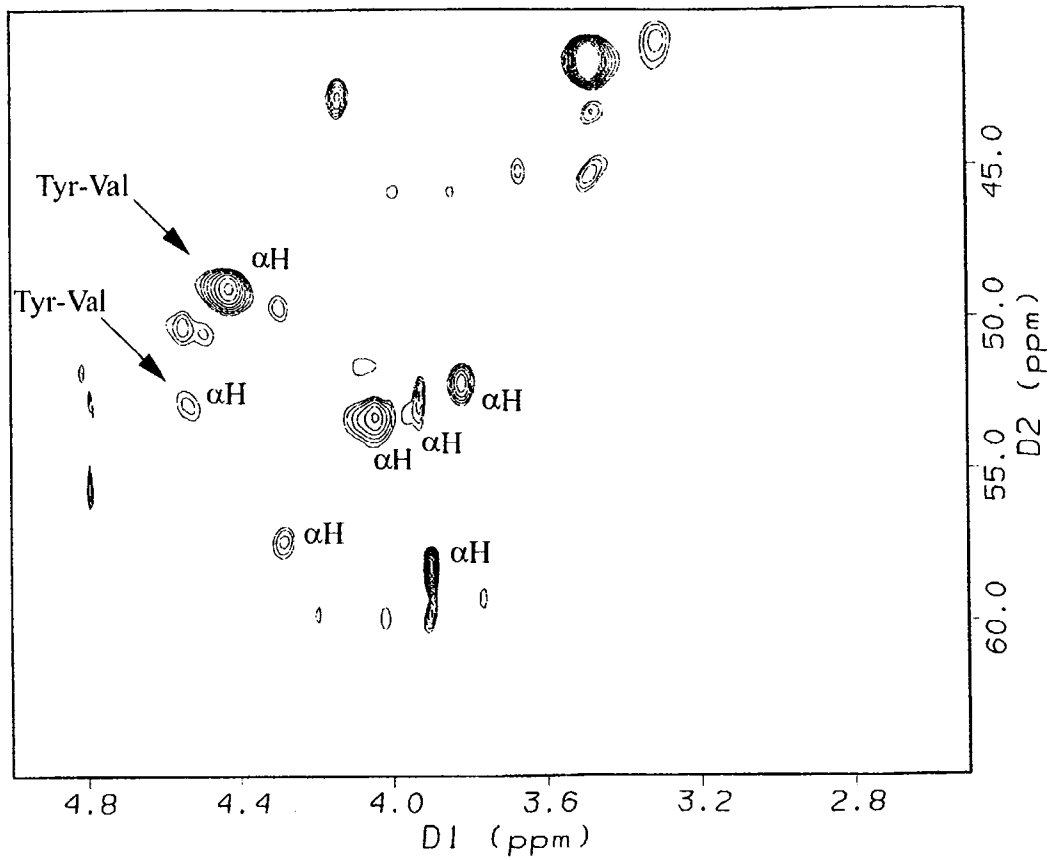


Figure III-9. INV4TP spectrum of ^{15}N Val $^{13}C_{\alpha}$ Tyr dPGK. The seven crosspeaks labeled αH arise from interactions between the $^{13}C_{\alpha}$ of tyrosine and its carbon bound proton. The crosspeaks identified in the HNCA experiment are also labeled. The D1 axis corresponds to the proton dimension and the D2 axis correpsonds to the carbon dimension. Protein concentration was 0.5 mM.

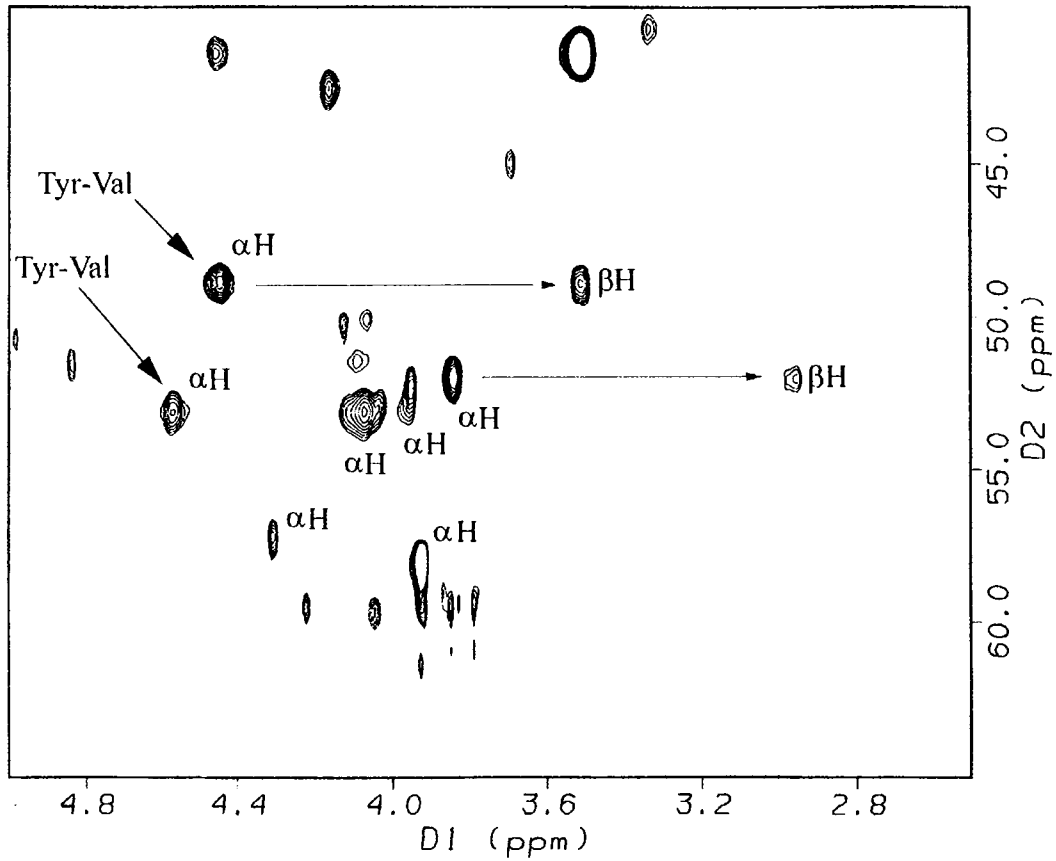
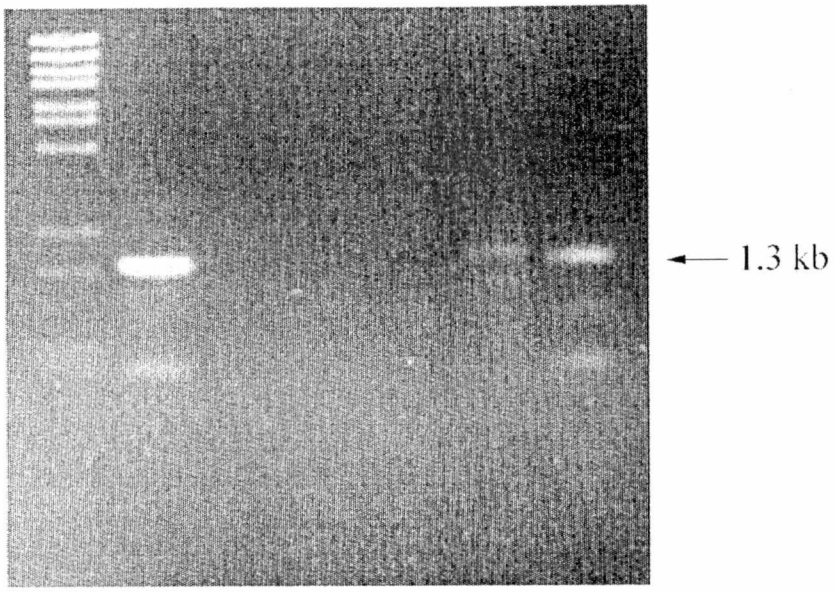


Figure III-10. INV4MLND spectrum of ^{15}N Val $^{13}C_{\alpha}$ Tyr dPGK. The correlation of the H_{α}'s and H_{β}'s of tyrosine are indicated by arrows. Both the H_{α}'s and H_{β}'s are labeled. The crosspeaks identified in the HNCA experiment are also labeled. The D1 axis correpsonds to the proton dimension and the D2 axis corresponds to the carbon dimension. Protein concentration was 0.5 mM.

*Nde*I Restriction Site: C A T A T G
Primer 1: T A T A A A A C A T A T G T C T T T A
yPGK: 281-T A T A A A A A C A A T G T C T T T A-301

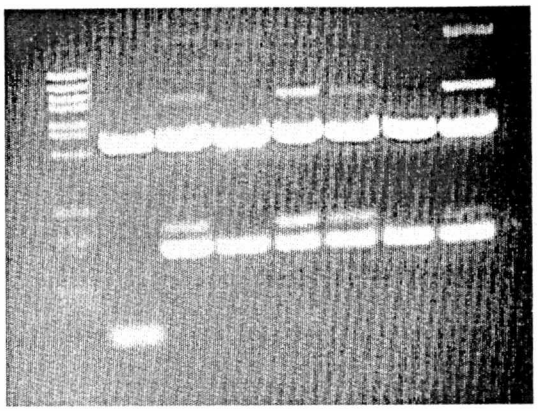
Primer 2: A T A A A T T C A T A T G A A T T G
yPGK 1539-A T A A A T T C A T A T G A A T T G-1557

Figure IV-1. Primers used in PCR of the yeast *PGK* gene. The introduced *Nde*I site is indicated. Conditions for PCR are described in Experimental Procedures, Chapter 4.

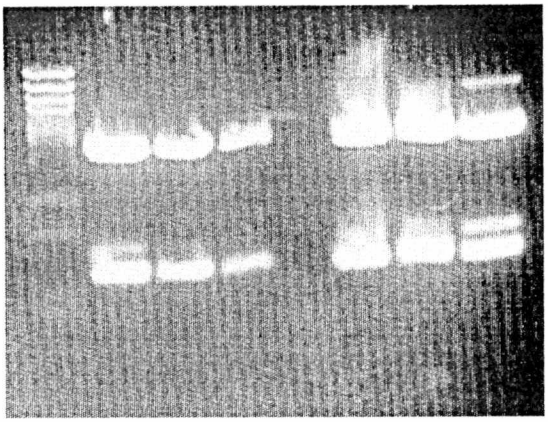


Lane 1: λ/*Bst*E Standards
Lane 2: Primer (25 pM), Template (250 pM), *Taq* polymerase (1 unit), Magnesium/Buffer (med), NTP's.
Lane 3: Template omitted
Lane 4: Primer omitted
Lane 5: *Taq* omitted
Lane 6: All reaction components (identical to reaction run in Lane 1)

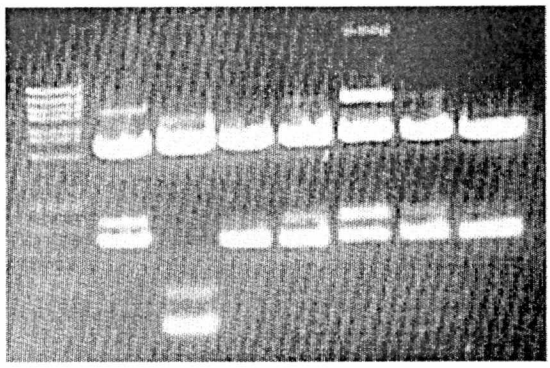
Figure IV-2. Gel showing the fragments obtained from PCR using the multicopy yeast plasmid *pCGY219* as the DNA template. The 1.3 kb fragment corresponding to PGK is indicated by the arrow.



Lane 1: λ/*Bst*E Standards
Lane 2: L12001
Lane 3: L12002
Lane 4: L1501
Lane 5: L1502
Lane 6: L22001
Lane 7: L22002
Lane 8: L2501

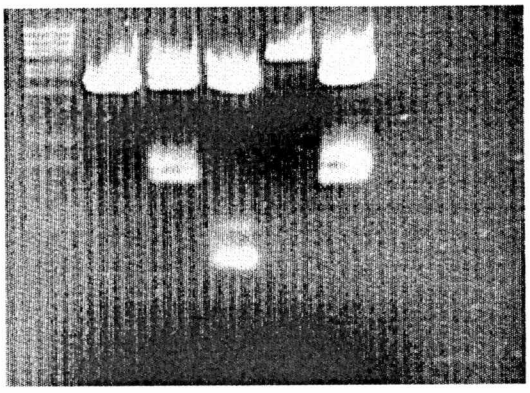


Lane 1: λ/*Bst*E Standards
Lane 2: L2502
Lane 3: L32001
Lane 4: L32002
Lane 5: L3501
Lane 6: L3502
Lane 7: L3503

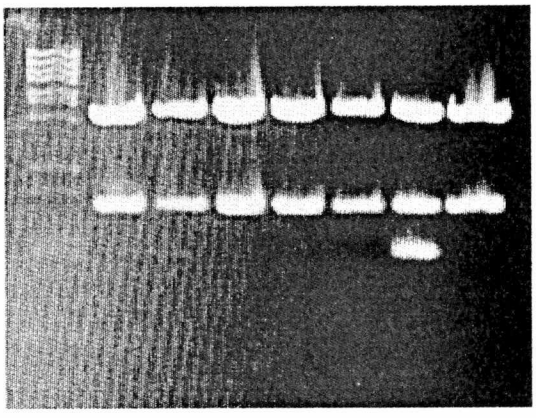


Lane 1: λ/*Bst*E Standards
Lane 2: L12003
Lane 3: L12004
Lane 4: L1503
Lane 5: L1504
Lane 6: L22003
Lane 7: L22004
Lane 8: L2503

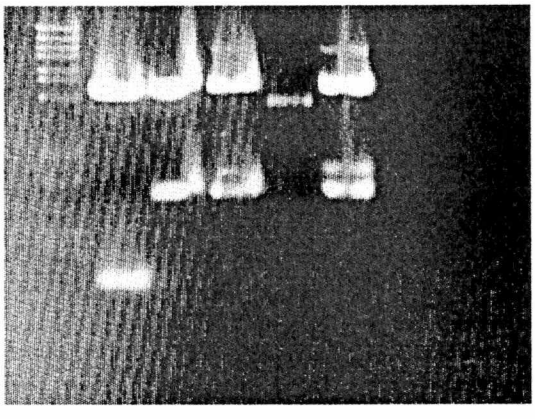
Figure IV-3. *Eco*RI digest of *pCRII/PGK* following ligation of the PCR product. The 1.3 kb fragment corresponding to yeast PGK is indicated. This digest confirms the presence of the PGK insert.



Lane 1: λ/*Bst*E Standards
Lane 2: L2504
Lane 3: L32002
Lane 4: L32004
Lane 5: L3503
Lane 6: l3504

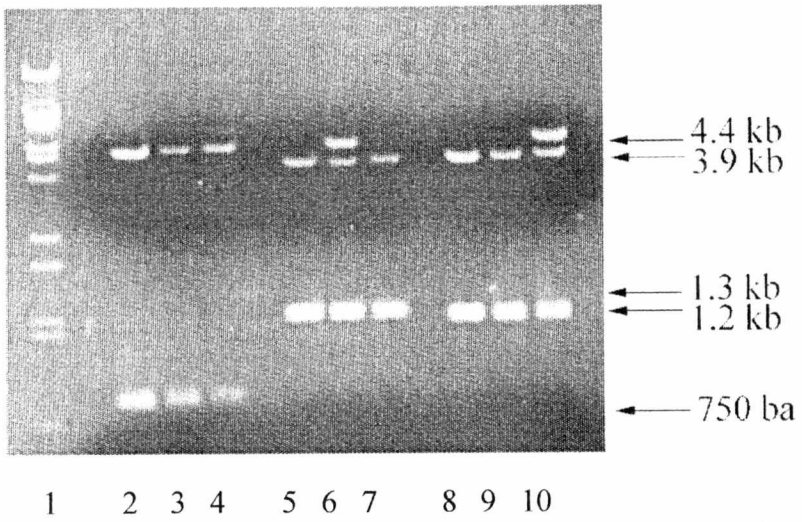


Lane 1: λ/*Bst*E Standards
Lane 2: L12005
Lane 3: L12006
Lane 4: L1505
Lane 5: L1506
Lane 6: L22005
Lane 7: L22006
Lane 8: L2505



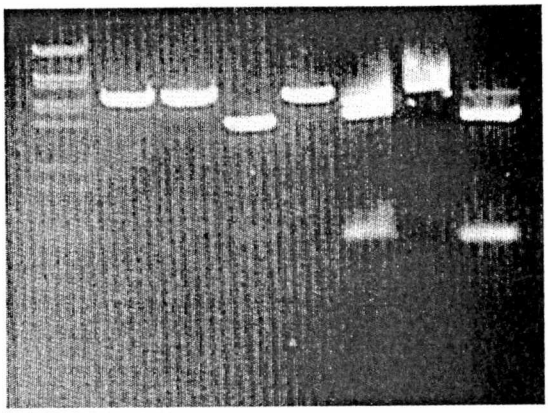
Lane 1: λ/*Bst*E Standards
Lane 2: L2506
Lane 3: L32005
Lane 4: L32006
Lane 5: L3505
Lane 6: L3506

Figure IV-3. (cont'd)

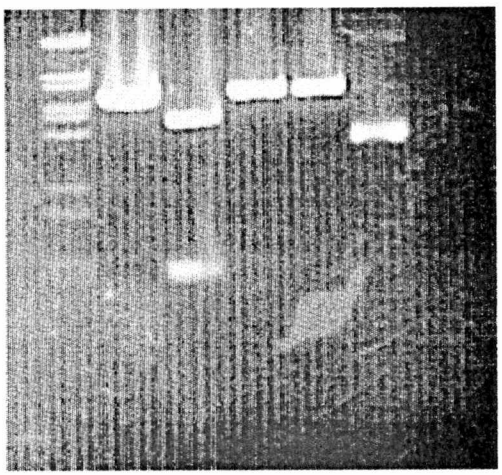


Lane 1: l/*Bst*E Standards
Lane 2: L6502 + *Kpn*I
Lane 3: L42002 + *Kpn*I
Lane 4: L62002 + *Kpn*I
Lane 5: L6502 + *Bam*HI/*Xho*I
Lane 6: L42002 + *Bam*HI/*Xho*I
Lane 7: L62002 + *Bam*HI/*Xho*I
Lane 8: L6502 + *Cla*I/*Not*I
Lane 9: L42002 + *Cla*I/*Not*I
Lane 10: L62002 + *Cla*I/*Not*I

Figure IV-4. Digests of *pCRII/yPGK* to determine the presence and orientation of the PGK insert. The *Kpn*I digest gave fragments of 750 ba and 4.4 kb, indicating that the insert is in the correct orientation in all three constructs. The *Bam*HI/*Xho*I digest gave fragments of 1.3 kb and 3.9 kb, indicating that the insert is present in all three constructs. The *Cla*I/*Not*I digest gave fragments of 1.2 kb and 3.9 kb, giving confirmation that the insert was in the correct orientation in all three constructs.



Lane 1: λ/*Bst*E Standards
Lane 2: L4501
Lane 3: L62003
Lane 4: L62001
Lane 5: L62004
Lane 6: L62002
Lane 7: L42001
Lane 8: L6502



Lane 1: λ/*Bst*E Standards
Lane 2: L6501
Lane 3: L42002
Lane 4: L42002 (no NdeI)
Lane 5: L42002 (no BamHI)
Lane 6: L42002 (no REN's)

Figure IV-5. Digest of *pCRII/PGK* constructs to confirm the presence of the required *Nde*I and *Bam*HI restriction sites. These sites flank the insert. The 1.3 kb fragment corresponding to yeast PGK is indicated.

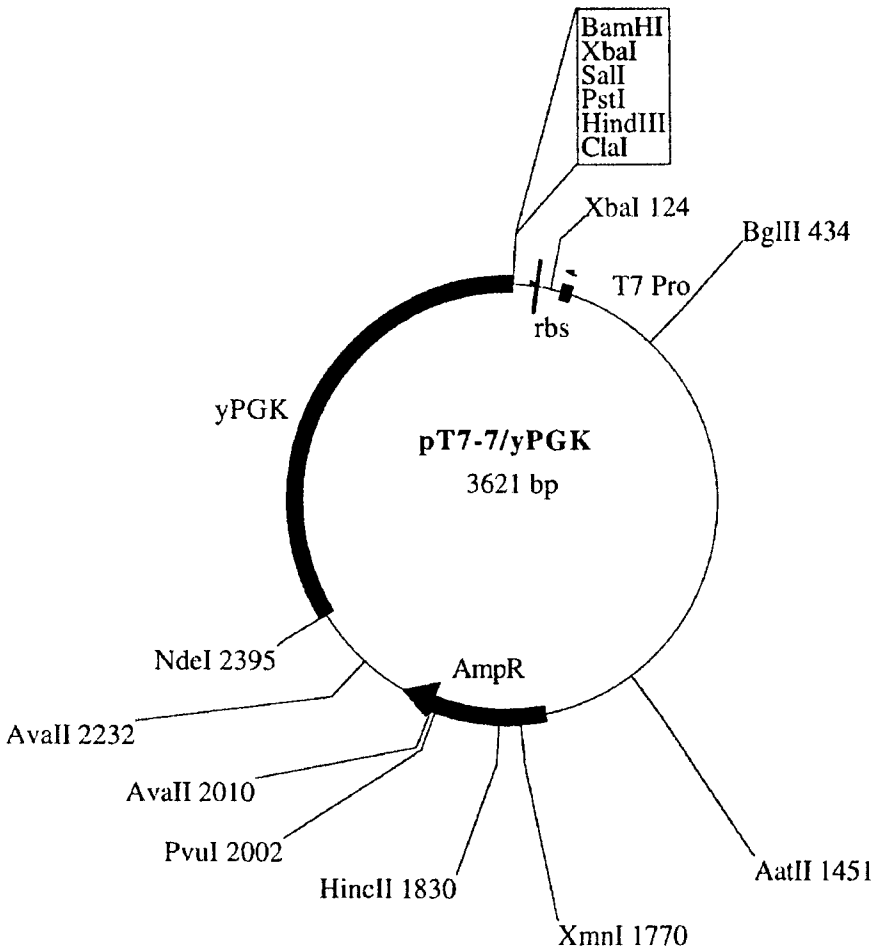
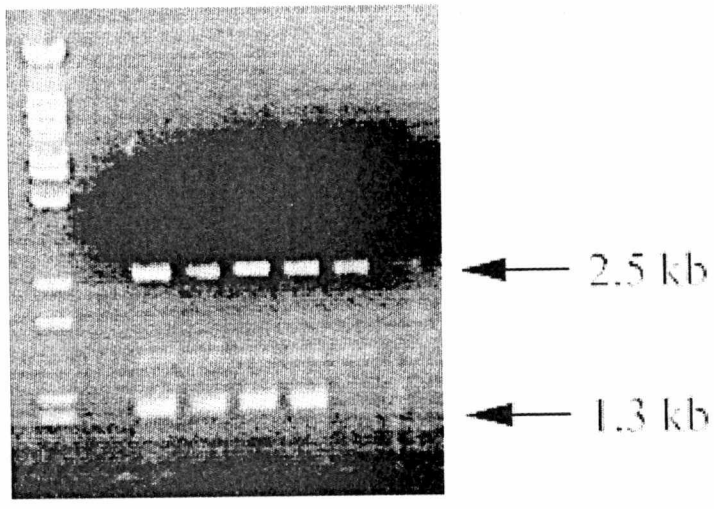
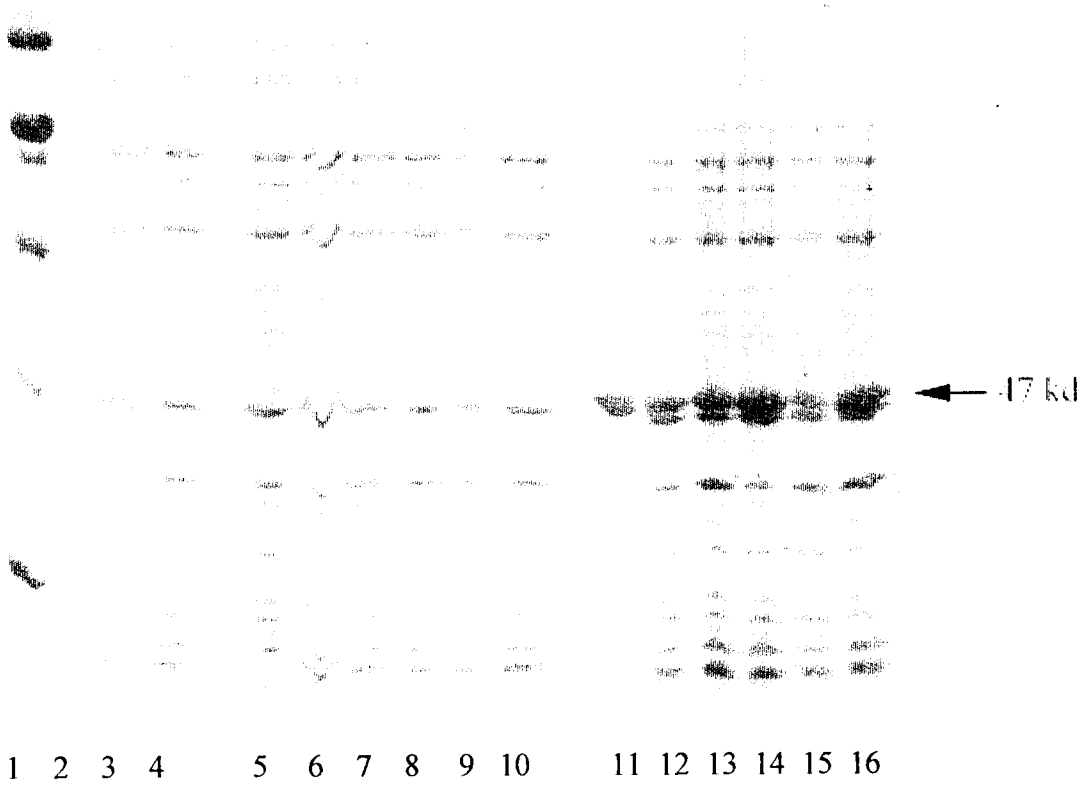


Figure IV-6. Plasmid map of the *pT7-7/yPGK* construct. The T7 promoter and ampicillin resistance sites are indicated, as are the *Nde*I and *Bam*HI sites used to ligate the PGK insert into the vector.



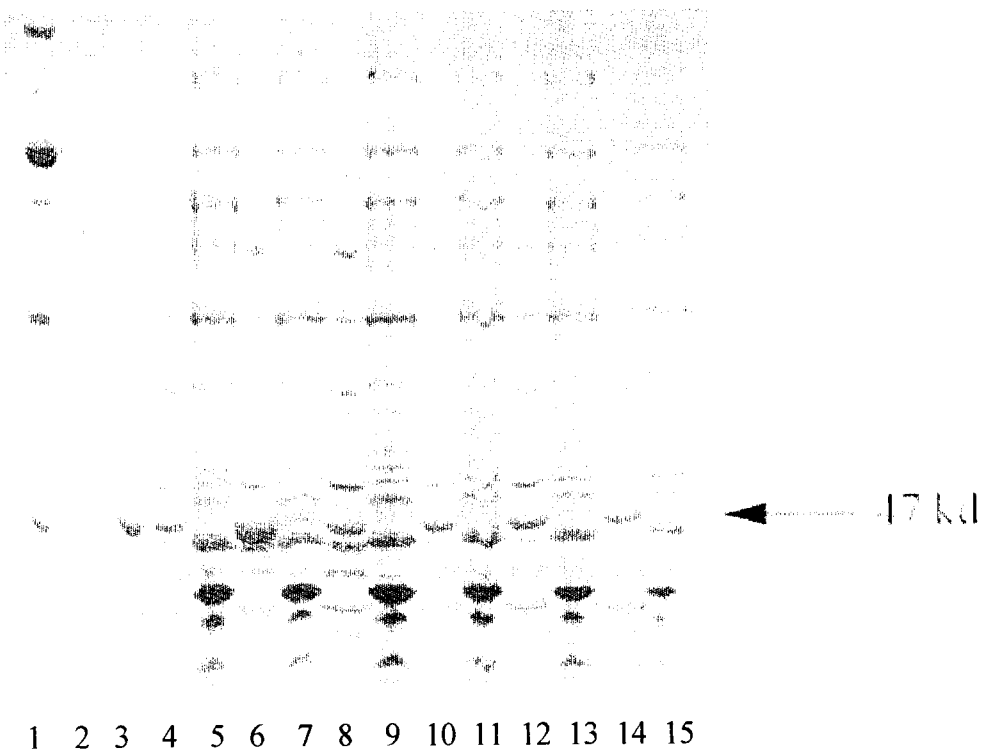
Lane 1: λ/*Bst*E Standards
Lane 2: L420023λ2
Lane 3: L420023λ1
Lane 4: L65023λ2
Lane 5: L65023λ1
Lane 6: L420023λ1
Lane 7: *pT7-7* (digested)

Figure IV-7. *Nde*I/*Bam*HI digest of *pT7-7/yPGK* ligation products. The constructs L42002 and L6502 were isolated from trasformed DH5α cells. The 1.3 kb PGK fragment and the 2.5 kb *pT7-7* fragments are indicated.



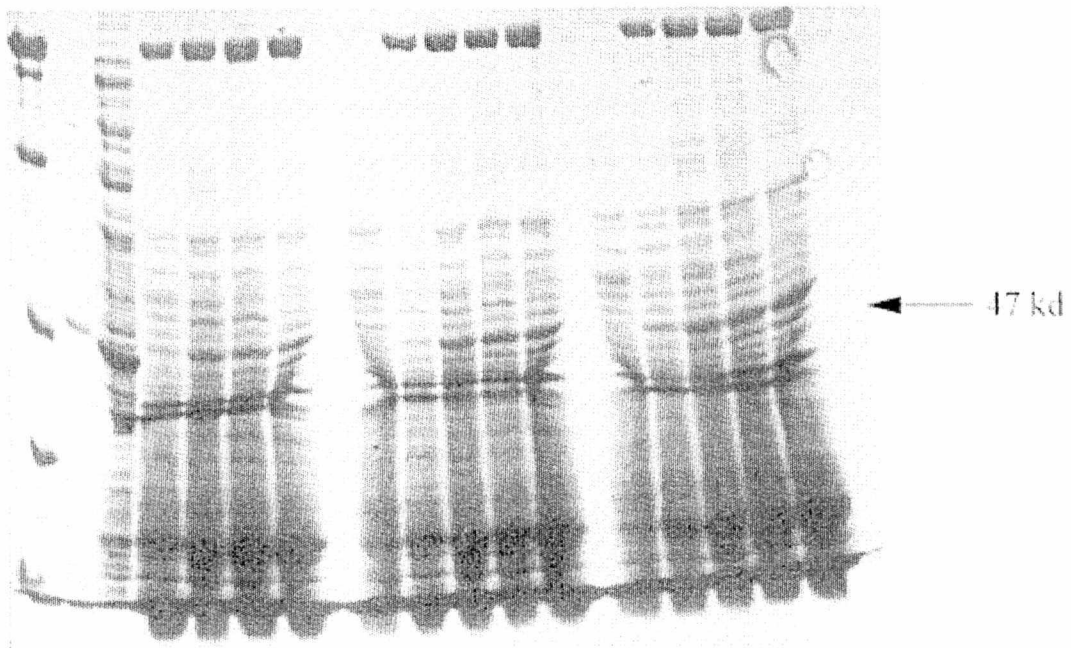
Lane 1: Molecular Weight Standards
Lane 2: PGK Standard
Lane 3: *pT7-7* (no insert)
Lane 4: *pT7-7* (no insert)
Lane 5: L420021 (uninduced)
Lane 6: L420022 (uninduced)
Lane 7: L420023 (uninduced)
Lane 8: L65021 (uninduced)
Lane 9: L65022 (uninduced)
Lane 10: L65023 (uninduced)
Lane 11: L420021 (induced)
Lane 12: L420022 (induced)
Lane 13: L420023 (induced)
Lane 14: L65021 (induced)
Lane 15: L65022 (induced)
Lane 16: L65023 (induced)

Figure IV-8. Protein gel showing the effect of IPTG induction on BL21(DE3)pLysS cells containing the constructs L6502 and L42002. These cells were induced for 12 hours and lysed as described in Experimental Procedures, Chapter 4. The 47 kDa band corresponding to yeast PGK is indicated.



Lane 1: Molecular Weight Standards
Lane 2: Lysozyme Standard
Lane 3: PGK Standard
Lane 4: L65021 Supernatant
Lane 5: L65021 Pellet
Lane 6: L65022 Supernatant
Lane 7: L65022 Pellet
Lane 8: L65023 Supernatant
Lane 9: L65023 Pellet
Lane 10: L420021 Supernatant
Lane 11: L420021 Pellet
Lane 12: L420022 Supernatant
Lane 13: L420022 Pellet
Lane 14: L420023 Supernatant
Lane 15: L420023 Pellet

Figure IV-9. Protein gel showing the presence of yeast PGK in the soluble fraction of induced BL21(DE3)pLysS cells containing the constructs L6502 and L42002. Cells were induced overnight and lysed as described in Experimental Procedures, Chapter 4.



Lane 1: Molecular Weight Standards
Lane 2: PGK Standard
Lane 3: L65021, t = 0
Lane 4: L65021, t = 1 hour
Lane 5: L65021, t = 2 hours
Lane 6: L65021, t = 3 hours
Lane 7: L65021, t = overnight
Lane 8: L65022, t = 0
Lane 9: L65022, t = 1 hour
Lane 10: L65022, t = 2 hours
Lane 11: L65022, t = 3 hours
Lane 12: L65022, t = overnight
Lane 13: L42002, t = 0
Lane 14: L42002, t = 1 hour
Lane 15: L42002, t = 2 hours
Lane 16: L42002, t = 3 hours
Lane 17: L42002, t = overnight

Figure IV-10. Protein gel showing the effect of different induction times on the production of yeast PGK. BL21(DE3)pLysS cells containing the constructs L6502 and L42002 were used in this experiment. Cells were lysed as described in Experimental Procedures, Chapter 4. The 47 kDa band corresponding to yeast PGK is indicated.

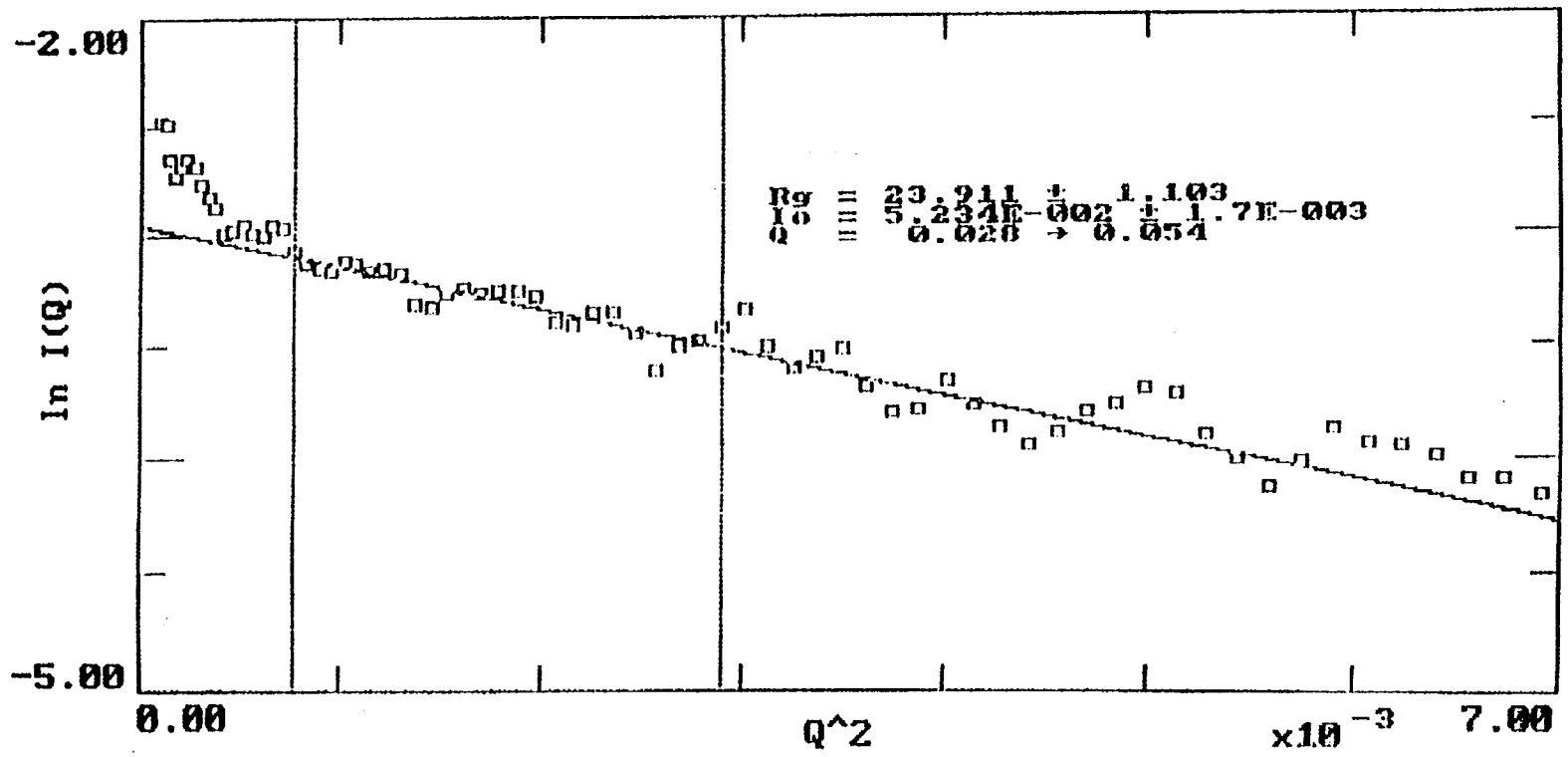


Figure V-1. Representative Guinier plot of SANS data acquired at ORNL. The concentration of dPGK in this sample was 4.8 mg/mL. The R_G of this sample was found to be 23.9 Å, which is in good agreement with previous experiments on isotopically normal PGK, which has an R_G of 24.0 Å. Sample concentration was 4.8 mg/mL. I(Q) is the intensity of scattering as a function of angle. Q^2 describes the angle of scattering.

Appendix B

Tables

Table V-1. Protein Concentrations And Enzyme Ratios Used in ORNL SANS Experiments

	PGK	GAPDH	Ratio (PGK:GAPDH)
Experiment 1	698 μM	171 μM	4:1
Experiment 2	375 μM	125 μM	3:1
Experiment 3	250 μM	125 μM	2:1
Experiment 4	163 μM	163 μM	1:1
Experiment 5	80 μM	160 μM	1:2
Experiment 6	80 μM	320 μM	1:4

Vita

John Allan Lamerdin was born on September 3, 1966 in San Francisco, California. He was raised and attended school in the Bay Area. In 1976, he and his family were fortunate to have an opportunity to live overseas in Tehran, Iran for about two and a half years. He returned to the San Francisco area and graduated from Novato High School in 1984. He did his undergraduate studies at The University of Redlands in Redlands, California, graduating in 1988 with degrees in both chemistry and philosophy.

John continued decided to pursue his studies in biochemistry at the graduate level and in 1988, he entered the Department of Chemistry and Biochemistry at San Francisco State University. He graduated in 1992 with a Master's degree in biochemistry. During this time he also worked as a formulation chemist at a research & development laboratory, as well as an oenologist at a local winery.

In August 1992, John joined the Department of Biochemistry at the University of Tennessee, Knoxville as a PhD candidate. During his time at UTK, John had the opportunity to present his research in the form of papers as well as at several different national meetings. He was awarded his PhD in August, 1997 and is pursuing a career in biotechnology patent law.