Windowed Direct Exponential Curve Resolution Quantification of Nuclear Magnetic Resonance Spectroscopy with Applications to Amniotic Fluid Metabonomics

Lucy L. Botros

Department of Chemistry

McGill University, Montreal

February 2007

A thesis submitted to McGill University

in partial fulfillment of the requirements of the degree of

Master of Science

Your file Votre référence
ISBN: 978-0-494-32665-7
Our file Notre référence
ISBN: 978-0-494-32665-7

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would first like to acknowledge my supervisor, Dr. David H. Burns for his support, guidance, incredible enthusiasm and seemingly endless flow of thought and ideas. It has truly been a pleasure to work for you.

I would next like to thank the entire Burns lab research group: Jon, Fabiano, Shing, Scott, David, Francis, and Kristin. You have all experienced the ups and downs of research along side with me. You've made it all enjoyable and bearable. Thank you for the years of fun.

I would like to acknowledge Dr. Kristine G. Koski and the members of the Koski lab at the School of Dietetics and Human Nutrition at McGill University for supplying the amniotic fluid samples as part of the larger ongoing amniotic fluid study.

A sincere thank you to Dr. Tara Sprules of the Quebec/Eastern Canada High Field NMR Facility, located at McGill University, for her time and the use of the 500MHz NMR Varian Spectrometer. Also to Dr. Paul Xia of the McGill Chemistry NMR facility for his help in setting up the excitation sculpting experiment.

Finally, I'd like to acknowledge my family and friends who've listened to me complain and rejoice about my experiences over these past years. Thank you for never leaving my side. I love you all.

ABSTRACT

This thesis presents a quantitative protocol of proton nuclear magnetic resonance (^{1}H NMR) that allows the determination of human amniotic fluid metabolite concentrations, which are then used in a metabonomic study to establish patient health during gestation. ^{1}H NMR free inductive decays (FIDs) of 258 human amniotic fluid samples from a 500MHz spectrometer are acquired. Quantitative analyses methods in both the frequency- and time-domain are carried out and compared. Frequency-domain analysis is accomplished by integration of the metabolite peaks before and after the inclusion of a known standard addition of alanine. Time-domain analysis is accomplished by the direct exponential curve resolution algorithm (DECRA). Both techniques are assessed by applications to calibration biological solutions and a simulated data set. The DECRA method proves to be a more accurate and precise route for quantitative analysis, and is included in the developed protocol. Well-defined peaks of various components are visible in the frequency-domain ^{1}H NMR spectra, including lactate, alanine, acetate, citrate, choline, glycine, and glucose. All are quantified with the proposed protocol. Statistical t-test and notched box and whisker plots are used to compare means of metabolite concentrations for diabetic and normal patients. Glucose, glycine, and choline are all found to correlate with gestational diabetes mellitus early in gestation. With further development, time-domain quantitative ^{1}H NMR has potential to become a robust diagnostic tool for gestational health.

RÉSUMÉ

La thèse présente un protocole quantitatif de la résonance magnétique nucléaire de proton (^{1}H RMN) qui permet la détermination des concentrations métabolite de fluide amniotique humaines, qui sont alors employées dans une étude metabonomique pour établir la santé patiente pendant la gestation. ^{1}H RMN « free inductive decay» (FIDs) sont collectés pour 258 échantillons de fluide humain amniotique provenant d'un spectromètre 500MHz. Des méthodes d'analyse quantitatives dans la fréquence et le temps-domaine sont effectuées et comparées. L'analyse frequence-domaine est accomplie par l'intégration des sommets de métabolite avant et après l'inclusion d'une addition standard connue d'alanine. L'analyse temps-domaine est accomplie par la « direct exponential curve resolution algorithm » (DECRA). Les deux techniques sont évaluées par des applications des solutions standards de calibration et de données simulées. La méthode de DECRA est un procédé beaucoup plus exact et précis pour l'analyse quantitative et est incluse dans le protocole développé. Des sommets bien definis de composante variées sont évidentes dans les spectres ^{1}H RMN du frequence-domaine, y compris le lactate, l'alanine, l'acétate, le citrate, la choline, la glycine, et le glucose. Tous sont quantifies avec le protocole propose. L'essai statistique, test T, et des diagrammes de type box-plot (boîte à moustaches) sont utilisés comme moyen de comparer des moyens des concentrations métabolite pour les patients diabétiques et non-diabétiques. Le glucose, la glycine et la choline se trouvent toutes à être en correlation avec le diabète mellitus gestationnel. Avec plus de developpement, le temps-domaine quantitatif ^{1}H RMN et ses applications aux mesures *in vivo* ont le potentiel de devenir un outil robuste pour la santé gestationnelle.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

NMR – Nuclear Magnetic Resonance

MRS – Magnetic Resonance Spectroscopy

MRI – Magnetic Resonance Imaging

RF – Radio frequency

FID – Free Inductive Decay

sw – Spectral Width

at – Acquisition Time

pw – Pulse Width

dt – Delay Time

S/N – Signal to noise ratio

nt – Number of FIDs accumulated and summed

GRAM – Generalized Rank Annihilation Method

DECRA – Direct Exponential Curve Resolution Algorithm

AF – Amniotic Fluid

GDM – Gestational Diabetes Mellitus

IDDM – Insulin Dependent Diabetes Mellitus

TSP – sodium 3-(trimethyl-2,2,3,3)-1-propionic acid-d_4

SER – Standard Error about Regression

SEP – Standard Error of Prediction

C.V. – Coefficient of Variance

FSE – Fast Spin Echo

I. Introduction

The goal of this thesis is to develop a quantitative nuclear magnetic resonance (NMR) spectroscopic approach for biological fluid analysis to obtain metabolic signatures underlying disease. This broad goal incorporates both development of magnetic resonance techniques and application of these techniques to a target biological system. In this introductory chapter, the fundamental principle in obtaining metabolic signatures will be discussed. Specifically, the concept of metabonomics is presented. Chapter overviews composing the thesis will be given at the close of this introduction.

Metabonomics

Metabonomics is the quantitative study of metabolic responses of living systems to pathological and physiological stimuli or genetic modification. Chemical profiles of metabolism as a response to disease are fingerprints for the processes (or the lack of processes) that are occurring simultaneously. Metabonomics is successful because these abnormalities cause perturbations of metabolite concentrations involved in metabolism and cellular pathways. Relative and absolute analysis of these metabolite concentrations can indirectly detect diseases in a patient and is implemented routinely in the medical community to decipher clinical health.

The application of metabonomics involves the analysis of multiple biological samples by a variation of characterization tools. Such biological fluids include those obtained non-invasively (e.g. urine), relatively easily (e.g. blood serum), and other fluids ranging from cerebrospinal fluid to seminal fluid [1]. Techniques such as infrared spectroscopy, Raman spectroscopy, and combinations of chromatography and mass

spectrometry are often applied in metabonomic studies. However nuclear magnetic resonance (NMR) spectroscopy has proved to be one of the most powerful technologies for biological fluid analysis and is capable of studying intact tissues, producing a comprehensive profile of metabolite signals, without the need of separation procedures [1].

Nuclear Magnetic Resonance Spectroscopy

NMR spectroscopy is a widely used characterization tool for identification and structural determination of many compounds. The spectrometers ranges from high to low resolution NMR (i.e. high to low field spectrometers), applicable to samples in both the liquid and solid state, and can be used in the detection of a wide range of nuclei, including ^{1}H, ^{13}C, ^{15}N, ^{19}F, and ^{31}P. NMR is capable of analyzing samples containing few simple compounds to a complex mixture of proteins and biological fluids. Yet, as broad as the qualitative NMR applications to chemistry are, there remains no clearly outlined protocol in the method of accurate quantitative analysis.

NMR has fundamental applications as a quantitative analytical technique. These applications have the potential to override commonly used methods for quantitative analysis today, if an accurate method were available. However, an accurate method for quantifying solutions of complex matrices at higher spectrometer resolutions has yet to be observed because of instrumental limitations. Current methods on simple solutions are limited to errors > 10% [2]. Whereas more accurate methods developed have been limited by low spectrometer resolutions [3]. For this reason, the first major purpose of this thesis is to further develop a quantitative protocol for high-resolution proton NMR (^{1}H NMR) of complex mixtures, specifically, biological fluids.

Application to Biological Fluids

Along with method development of ^{1}H NMR quantification, a metabonomic application is included as co-objective of this thesis. One biological fluid that is an ideal candidate for this metabonomic study using ^{1}H NMR, is amniotic fluid. Amniotic fluid contains a wide range of constituents that result from metabolic processes occurring throughout gestation. During pregnancy, regular quantification of these metabolites ensures the maturity progress of the fetus and the clinical health of the mother. Multiple individual tests are available and are the status quo for metabonomics of amniotic fluid today. However, applying ^{1}H NMR spectroscopy to amniotic fluid has an advantage in that the spectra produced are metabolic profiles, whose multiple signals can be analyzed simultaneously. Hence, the second major purpose of this thesis is to perform a metabonomic study on amniotic fluid using the developed quantitative ^{1}H NMR spectroscopy. This study will in effect map metabolic profiles of maternal and/or fetal gestational diseases and conditions.

Noting the availability of Magnetic Resonance Spectroscopy (MRS) and Magnetic Resonance Imaging (MRI), the application of ^{1}H NMR spectroscopy to biological fluid metabonomics is advantageous as a non-invasive diagnostic tool. Current individual metabolic tests incorporate the collection of amniotic fluid by an invasive method, amniocentesis, which has many associated complications and risks. The replacement of amniocentesis is the driving force in many current studies and research investigations. With mapped metabolic profiles of specific gestational diseases by high-resolution ^{1}H NMR, further prenatal diagnostic *in vivo* applications of this tool are possible.

Thesis Overview

The thesis is divided into four sections. First, the concept and theories of quantitative ^{1}H NMR are introduced. This introduction lists NMR's associated advantages over other spectroscopic techniques and quantitative methods. Theory presented describes the physical phenomenon that is magnetic resonance, the manner in which an NMR signal is generated, and the relationship between sample concentration and the NMR signal. Acquisition and post-acquisition parameter requirements for accurate quantitative measurements are also described. The chapter then goes on to describe past and present methods used in quantifying the NMR signal. Both frequency- and time-domain signal analyses are discussed. A critical review of previous quantitative NMR studies and the challenges that arise when making a quantitative measurement with NMR concludes the discussion.

An overview of the investigative approach is summarized in the third chapter. The development of the quantitative ^{1}H NMR protocol is detailed. Specifically, the experimental procedure used in the collection of NMR spectra for a set of calibration biological solutions is described. Two methods of quantification are applied to the above data. The first is a 'classical' integration method of specific proton peaks in the frequency-domain NMR spectra, before and after the inclusion of a known concentration standard addition. The second is a time-domain analysis method based on generalized rank annihilation. The direct exponential curve resolution algorithm is derived and applications to both simulated free inductive decays and calibration biological solutions follow.

In chapter IV, a review of the specific biological fluid to be studied, and a review of a gestational metabolic disorder are discussed. The chapter focuses on amniotic fluid and the development of gestational diabetes mellitus. Amniotic fluid's origin throughout gestation is specified along with the changing biochemical makeup. Also discussed are the current procedures for prenatal diagnosis. Focus is placed on amniocentesis, the protocol described, along with the associated problems and risks. Background information is supplied on gestational diabetes mellitus with a description of common biochemical tests used for screening this disorder. Specifically, the relationship of glucose and other metabolite concentrations are related to the diagnosis of gestational diabetes mellitus.

Finally, chapter V demonstrates the capability to statistically identify patients with gestational diabetes mellitus from healthy patients by quantitatively analyzing amniotic fluid ^{1}H NMR spectra using the direct exponential curve resolution algorithm. Constituents that show significant differences in concentration between the two patient groups are identified. These conclusions are the result of t-test analyses of means and the use of notched box and whisker plots, whose details are discussed.

The thesis ends with a chapter of general conclusions, which summarizes the findings of the present research, and discusses limitations to the quantitative method introduced. A discussion on future work and ambitions derived from this study follows. Future objectives include the application of the defined quantitative protocol to other biological fluids.

II. Nuclear Magnetic Resonance As A Quantitative Tool

Nuclear Magnetic Resonance (NMR) spectroscopy is a technique that exploits the magnetic properties of nuclei. NMR has advanced applications in chemistry, biology and biochemistry and could be used to determine (1) the structures of unknown molecules, (2) the absolute stereochemical configuration of known molecules, (3) the modes of interaction of small molecules with larger ones, (4) the kinetics and thermodynamics of such interactions, and (5) the dynamics of molecules in supramolecular assemblies [4]. For example, NMR spectroscopy applications are currently focused on the determination of protein structures and monitoring of protein folding and denaturing processes [5-7].

NMR spectroscopy can fundamentally be applied as a quantitative analytical method [2]. Yet, this spectroscopic method is used far more often for structural elucidation and identification of unknown compounds. With such complex applications, NMR spectroscopy has yet to be routinely applied to the seemingly simple task of quantification. It is thought that NMR's quantitative analysis power is not only overshadowed by its qualitative applications, but is greatly underestimated. Early reports regarding the achievable precision of quantification are not only inconsistent, but also tend to rebuff NMR as a precision method by estimating the error to be in the 10% range [2].

This chapter discusses previous quantitative NMR investigations and their achieved accuracy. Before the limitations to this accuracy are understood, it is necessary to comprehend the basics of NMR spectroscopy. The physical phenomenon of nuclear magnetic resonance is described, as well as how an NMR signal is generated. The fundamental relationship that dictates the quantitative nature of the NMR signal is then

presented. Next, NMR instrumental parameter optimizations, to ensure accurate quantitative analysis, are described. Finally, a literature review of common frequency-domain and time-domain quantitative analysis methods is presented, followed by advantages of these types of analyses.

2.1 Nuclear Magnetic Resonance Spectroscopy Basics

NMR is a physical phenomenon based upon the magnetic property of an atom's nucleus. All nuclei that contain a half integer nuclear spin (ex: protons have a nuclear spin of –1/2 or +1/2) have an intrinsic magnetic moment. When these nuclei are placed in an external magnetic field, their magnetic moments either align parallel or anti-parallel with this magnetic field, depending on the nuclear spin state (high and low energy states). As there is a slight difference in the population in these high and low energy states, an overall magnetization vector results. Note that the stronger the external magnetic field strength, the greater the difference in population between the lower and higher energy states, resulting in a larger overall magnetization vector.

Quantum mechanics requires that the magnetic moments of these nuclei not be aligned exactly parallel or antiparallel to the external magnetic field. Instead they are forced to remain at an angle to the external magnetic field, causing them to 'wobble' around the axis of the field at a fixed frequency [4]. This wobble is called 'precession'. These nuclear magnetic moments in a magnetic field also precess with a characteristic angular frequency called the Larmor frequency, which is a function of the external magnetic field strength. Depending on the surrounding chemical environment (i.e. neighboring nuclei, aromatic systems etc.), some nuclei are precessing at either slightly slower or faster frequencies than the Larmor frequency [4]. This is because the variable

chemical environments cause nuclei to experience slightly variable magnetic field gradients causing them to precess at slightly different frequencies. However, the resulting overall magnetization vector remains in precession around the external magnetic field axis with the Larmor frequency.

Signal Generation

The NMR signal is generated by irradiating the nuclei in the external magnetic field with a second brief but powerful radio-frequency (RF) pulse, perpendicular to the first external magnetic field. This transient RF pulse has a magnitude equal to that of the Larmor frequency [4]. The irradiation causes energy flips of the nuclei and tips the overall magnetization vector at an angle, θ, to the external magnetic field. The relaxation of the magnetization vector back to its original alignment with the external magnetic field (i.e. the process of the nuclei returning to the original population distribution prior to the RF pulse) is what is measured by the NMR spectrometer. This measured signal is defined as the time-domain NMR signal and is referred to as the free inductive decay (FID). An FID has an exponentially decaying character and is typically Fourier-transformed to produce the recognizable frequency-domain spectrum.

Fundamental Quantitative Relationship

The fundamental relationship that governs NMR's quantitative definition is that the resulting NMR signal is proportional to the number of nuclei in a given sample, as given by Equation 2.1.

$$I_X = N_X K_S \qquad 2.1$$

Where I_x is the NMR signal intensity for component x, N_x is the number of nuclei of component x within the sample, and K_S is a proportionality constant resulting from parameters of the spectrometer [8]. In other words the NMR signal is directly proportional to concentration of component x within the sample. Therefore if component x can be successfully resolved from the entire NMR signal, the amount within the sample can be theoretically computed.

Numerous studies have attempted to quantify NMR signals accurately by applying a wide variety of methods. Frequency-domain analyses are considered to be the 'classical' quantitative analysis route, using Equation 2.1 and peak integrals as the basis for all applications in this domain. Recent attempts to quantify the NMR signal in the time-domain, that are based on analysis of the resulting FID signal prior to Fourier transformation, have also been made. Independent of both frequency-domain and time-domain methods, the equality quoted in Equation 2.1 must be valid in order to obtain precise and accurate quantitative estimates. In other words, the NMR signal, I_x, must be proportional to the number of nuclei in a given sample. To ensure this, a number of instrumental parameters, prior to and after acquisition must be 'tuned' to proper settings to obtain quantifiable signals. The following section 2.2 lists these parameters with a brief description of each.

2.2 Parameters For Accurate Quantitative NMR Analysis

To ensure the intensity of a ^{1}H NMR signal is accurately and reproducibly proportional to the number of nuclei, a number of data acquisition and post-acquisition parameters need to be carefully optimized. These quantitative experimental conditions

include digital resolution (i.e. acquiring enough data points), adjusting the width of the 90° transient RF pulse (*pw*), adjusting relaxation delay time (*dt*), good S/N and peak shape, and proper data processing techniques [2]. Before one can properly optimize these experimental conditions, the experiment itself must be understood. In figure 2.1, the single-pulse NMR experiment is depicted.

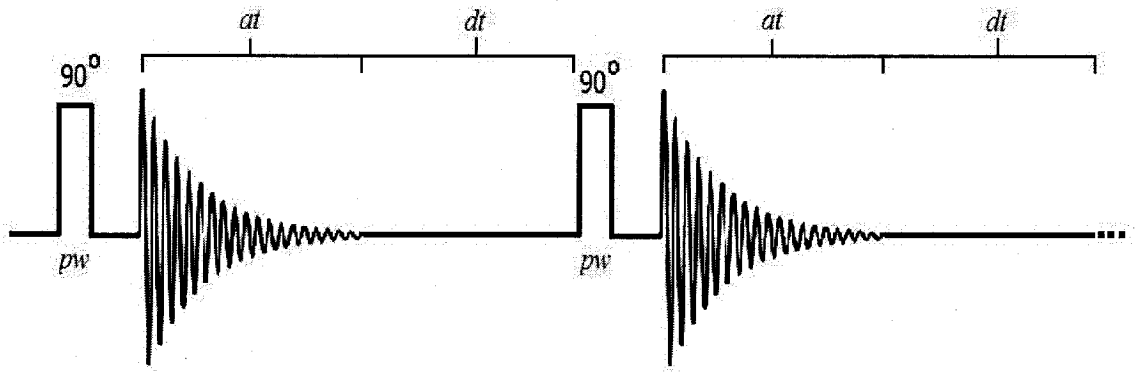

Figure 2.1 – Schematic representation of the single-pulse NMR experiment. Where *at* = acquisition time, *dt* = relaxation delay time, *pw* = 90° RF pulse width

The RF transient pulse (*pw*) is followed by a short dead time, after which data acquisition of the FID begins. Turning the receiver on initializes this acquisition, and the digitization of the FID is carried out by the analog-digital converter throughout the acquisition time, *at*. After a relaxation delay of *dt*, during which no acquisition occurs, the procedure is repeated, and the cycle continues number of times, *nt*, specified by the user. Fourier transformation of the time-averaged and digitized FID then yields the frequency-domain spectrum. The FID can be viewed as the superimposed sum of exponentially decaying sinusoids corresponding to the various nuclei in the sample, and an added modulation by random noise [9]. Now that a visualization of this experiment has been given, the optimization of these experimental parameters can be discussed

Digital Resolution

A pulse sequence for obtaining a ^{1}H NMR spectrum contains three basic parameters that are under user control: the relaxation delay, *dt*, the pulse width in μs, *pw*, and the acquisition time, *at*. The acquisition time must be set to a long enough time in order adhere to the Nyquist theorem, which is related ($2\times$) to the digital resolution of the signal. In other words, the longer the *at,* the more data points are used to define the signal. The length of *at* is generally selected on the basis of the spectral width (*sw*) and the required digital resolution for that *sw*. Spectral width, *sw,* is defined as the range of nuclear precession frequencies to be examined. For most ^{1}H NMR spectra, *at* is selected so that the spectral resolution (at^{-1}) is around 0.25 Hz or better [2]. Although *at* is set prior to acquisition of the NMR signal, improvement of the digital resolution can be further enhanced by post-acquisition data processing techniques such as zero-filling.

Pulse Width

The pulse width is the length in μs of the RF pulse (*pw,* 90_x°), centered at the spectrometer operating frequency (also known as the Larmor frequency), v_0, where all of the magnetization along the z-axis is converted into transverse (or x,y-) magnetization, as seen by the figure 2.2 below.

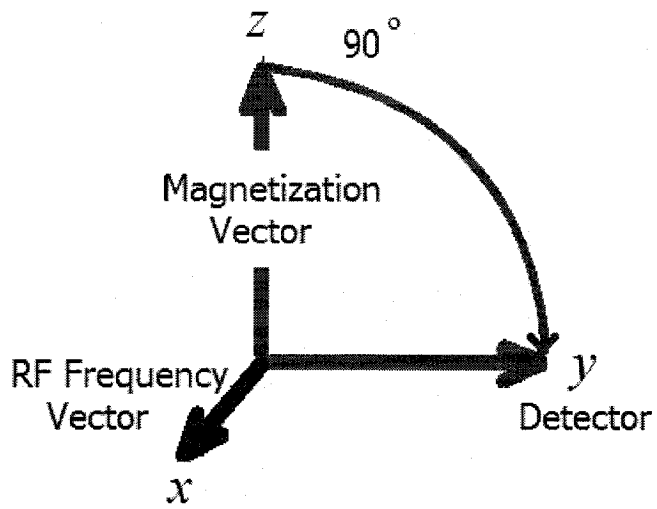

Figure 2.2 – Effect of a 90_x° RF pulse. The z-magnetization vector is rotated to the x,y-plane

Theoretically, if a 90° RF pulse is used, all of the z-magnetization is converted into the x,y-magnetization, along which the signal is detected, resulting in maximum sensitivity and signal intensity [2].

Not only must this RF pulse be 90°, it must also be sufficiently short enough to excite the entire spectral region uniformly [10]. As a direct consequence of the uncertainty principle, this brief pulse acts as if it covers a range of frequencies from $v_0 - \Delta v$ to $v_0 + \Delta v$, where Δv is inversely proportional to *pw*. Thus the shorter *pw*, the greater range of frequencies excited [4].

Relaxation Delay

The relaxation delay time (*dt*) represents the time, in seconds, for the equilibrium magnetization to be re-established between transient RF pulses. A long enough *dt* must be used in order to allow the equilibrium population distribution to be re-established before the next pulse can begin. The length of *dt* is governed by the relaxation properties

of the nuclei in the mixture and is characterized by their individual relaxation times [2]. Relaxation times fall into two types: longitudinal relaxation (T_1) and transverse relaxation (T_2). The latter of these two relaxation times is discussed in the following 'S/N and Peak Shape' section.

The first relaxation time, T_1, governs the time required for the collection of nuclei placed in an external magnetic field to establish a new (Boltzmann) equilibrium distribution with a slight excess of nuclei in the lower energy state, resulting in the positive z-magnetization (refer to figure 2.2) [4]. The values of T_1 range broadly depending on the type of nucleus, the location of the nucleus within a molecule, the physical state of the sample (liquid or solid), and the temperature. To measure the T_1 values of the sample nuclei, inversion recovery experiments are employed [4]. This experiment applies the following pulse sequence: $180_x^o - \tau - 90_x^o$. Where τ is the relaxation delay time (in seconds), varied at constant intervals. The first 180_x^o pulse brings the z-magnetization along to the negative z-axis. If after a specific τ, and the subsequent 90_x^o RF pulse, no signal is observed for particular nuclei (no magnetization along the x,y plane), the value of τ is equal to that of the nucleus' longitudinal relaxation time, T_1.

Once the T_1 values of all nuclei in a sample are accurately known, adjustment of the delay time, *dt*, may be established and optimized. The delay time is optimized to five times the longest relaxation time, T_1(max), as shown in Equation 2.2.

$$dt = 5 \times T_1(\max) \qquad 2.2$$

It is well known, that waiting a minimum of $5 \times T_1(\max)$ allows 99.3% of the slowest relaxing nucleus to completely equilibrate, resulting in this percentage of the total signal intensity.

S/N and Peak Shape

If the classical method of NMR quantification is to be employed, collection of proper frequency-domain NMR spectra must first be ensured. Good signal to noise (S/N) and Lorentzian peak shapes are essential signal characteristics to reduce error in integration.

Adjusting the external magnetic field homogeneity can optimize peak shape. Prior to acquiring an NMR signal, the homogeneity of the external magnetic field is regulated by passing extremely small currents through a series of shim coils located around the probe (the sample vessel positioned along the center (z-axis) of the external magnetic field). This process is termed "shimming" a magnet. A badly shimmed magnet can cause distortions in the line-shape of an NMR peak in the frequency-domain spectrum. These distortions can manifest themselves as tailings, or humps and even broadening of the peaks. The peak broadening is a consequence of the relationship between the transverse relaxation time or spin-spin relaxation, T_2 and inhomogeneities in the magnetic field. The T_2 is inversely proportional to the line-width at half-height ($w_{1/2}$): $T_2 \sim (w_{1/2})^{-1}$. Consequently, slight inhomogeneities in a magnetic field cause a change in the decay rate of the FID (T_2) and thus broaden the line-widths at half-height.

Recall that signal to noise ratios (S/N) can be enhanced by optimizing both the RF pulse width (*pw*) and the relaxation delay time (*dt*). Another common practice to

increase S/N is to accumulate and sum multiple FIDs [11, 12]. In other words, after one pulse cycle and an FID has been collected, the same relaxation delay time is set before the next pulse cycle begins and another FID collected. This cycle continues until the number of FIDs specified by the user, *nt*, (normally a multiple of four) is acquired and summed. The S/N enhancement is $(nt)^{1/2}$ [11, 13]. It has been stated that the detection limit for the lowest quantifiable concentration is 3×signal to baseline noise [13]. Another study discloses that a S/N of at least 150 is required for the target uncertainty of 1% [8].

Post-acquisition Processing

Once the NMR signal has been collected, the data can then processed. These post-acquisition processing techniques include zero-filling and spectral phasing. The process of zero-filling is done prior to Fourier transformation of the time-domain FID. This practice is when zeros are added to the end of the FID to increase the number of data points. Upon Fourier transformation, zero-filling improves the resolution of the frequency-domain spectrum. The "rule of thumb" is to always zero-fill so that the number of points in the Fourier transformed spectrum is a base power of 2 (ex: 2^{13}) [2]. It has been described, that at least five data points must appear above the line-width at half-height of each peak for a precise and reliable subsequent integration [8]. Meaning, this resolution enhancement consequently improves integration reproducibility in the classical method analysis.

Spectral phasing is a process where the baseline of the frequency-domain NMR spectrum is adjusted to maximize the baseline flatness. The NMR spectrum usually requires correction by a phase angle, θ. The phase angle contains both a frequency-

independent and a frequency-dependent component. The former arises during the signal detection procedure and can be corrected for by a zero-order phase correction. The latter arises from the necessary delay, termed the dead time, prior to acquisition of the NMR signal and can be corrected using a first-order phase correction [3]. Careful phasing of the spectrum ensures that the integrals representing the various signals in the Fourier transformed spectrum have minimal distortion and therefore, contributes to quantitative reproducibility. Both manual and automatic phase corrections are available. However, it is far more common to manual phase a spectrum over the automatic alternative, as the automatic routines may produce slightly wrong phase and baseline corrections [2, 8].

The present experiment was designed in order to obtain ^{1}H NMR signals of calibration biological solutions and amniotic fluid samples that adhere to the fundamental relationship given in Equation 2.1. Acquisition and post-acquisition parameters such as the relaxation delay times, S/N, and digital resolution are optimized, while still considering the duration of the entire experimental procedure. Because 250+ samples were collected and measured, it was necessary to limit the total acquisition time per sample, while ensuring quantitative ^{1}H NMR signals.

2.3 Current Methods for Quantitative NMR Analysis

Frequency-domain Analysis Methods

The most 'familiar' form of the NMR signal is the frequency-domain spectrum, where signal amplitude is plotted versus frequency. It is this NMR spectrum that is analyzed and used for spectral interpretation in terms of both structure and quantification. The fundamental relationship outlined in Equation 2.1, is applied to the frequency-

domain spectrum, where I_x represents the integration of the chosen region within the spectrum. It is common to compute nuclei area ratios (I_{x_1}/I_{x_2}) to obtain relative quantitative results, similar to an internal standard methodology. This technique has no need for an external calibration model or an intensity calibration model unless absolute concentration measures are desired. This is the most common method of analyzing an NMR spectrum quantitatively, as is referred to in this thesis as the 'classical' quantitative method.

An example of this nuclei area ratio technique is given by Jay L. Bock [14] , who instead of using signal integrals, determined the ratios of signal intensities when attempting to quantify significant metabolites in amniotic fluid, using a 600MHz NMR spectrometer. Because the absolute intensity scale varied between samples, he arbitrarily selected the citrate signal as an internal intensity standard. The other metabolite signal intensities were expressed as the ratios of their peak heights to that of the citrate peak. Linear correlation plots were constructed where known concentrations were plotted versus metabolite intensity ratios, and R^2 values of 0.93 and 0.97 were obtained. Holzgrabe et al [12] also applied this ratio method in an attempt to evaluate the quality of drugs. Levels of active ingredients and impurities in drugs were measured in order to fulfill requirements outlined by the European pharmacopoeia (PhEur 5.0) and the United States pharmacopoeia (USP 26).

This 'classical' integration method can be applied to varying samples with different properties. Integration analysis can be coupled to classical analytical methodologies in order to determine absolute concentrations. Methods such as standard additions and external standards have all been previously applied in many studies. The

first by McGowan et al [13, 15], in an attempt to quantify amniotic fluid, used a 400MHz spectrometer and standard addition methodology. Quantification of metabolites in this study was achieved by inserting three standard additions of known quantity of alanine directly into the analytical sample. The key metabolite signals were integrated before and after each standard addition. A calibration curve was setup and the initial quantity of alanine within the sample was extrapolated. By compensating for the equivalent number of nuclei, and the application of integral ratios, the other important species were quantified. The standard error of mean metabolite concentrations derived from this study was an average of 22.8%.

The second study applied an external standard methodology to quantify natural products and other low molecular weight compounds. Burton et al [16] placed equal volumes of standard and analyte solutions in separate but identical NMR tubes. Spectra of the standards and analytes were recorded under the same conditions and pulse sequences. Linear regressions of standard signal integration versus gravimetric concentration were constructed and showed "excellent linearity", with R^2 values ranging from 0.9982 – 0.9999. From these linear regressions, Burton et al were able to estimate unknown analyte concentrations with a reported precision and accuracy of 1%.

Many other studies have been reported using the classical integration method. Their experience shows that typical uncertainties with this technique fall within the 0.5-2% range. However, methods of integration can vary greatly between users resulting in the lack of reproducibility between laboratories. One study even reported up to 100% difference in results between over 30 participating university laboratories, research institutes, and companies [8]. Because of this user dependence, other techniques have

surfaced in the attempt to properly integrate a signal in a frequency-domain NMR spectrum.

These techniques include model functions and curve-fittings. An automated algorithm, developed by Crockford et al [17], used reference compound spectra to find best-fit transformations between signals in the reference compound and the corresponding signals in analyte spectra. By creating a database of replicate NMR measurements of biologically important compounds at various concentrations, Crockford et al were able to quantify these key compounds in biological fluids, such as urine. Other fitting techniques have also used linear combinations of reference spectra of known compounds. These best-fits were often obtained using a least-squares based method [18]. Other least-squares-based methods have made use of complex linear combinations of Lorentzian and Gaussian models and have been applied to the analysis of simple mixtures of organic compounds [19]. These techniques, however useful, are computationally tedious and decrease in accuracy when unique line-shapes do not exist, and can be distorted from poor signal acquisition and chemically induced changes.

Regardless of these innovative curve-fitting methods, the major source of error remains in the integration of signals in a frequency-domain NMR spectrum. There are a multitude of factors that can degrade the accuracy and precision of integration including user dependence, but also spectral and baseline distortions. To remedy this, the new common 'trend' in quantitative NMR analysis is to directly analyze the time-domain signal, the free inductive decay (FID).

Time-domain Analysis Methods

As opposed to frequency-domain quantitative methods, time-domain techniques analyze the FID directly without Fourier transformation into the classical and recognizable NMR spectrum. Nonetheless, the FID is still generated in the same manner by pulsed RFs exciting the nuclei in a given sample, resulting in a spectrum of signal amplitude versus time. The fundamental relationship in Equation 2.1 remains, and the resolvable components are theoretically quantifiable.

As it is visually simpler to quantify NMR signals in the frequency domain, it is more inherent to want to process the data in the domain it is collected. Many post-acquisition processes are also avoided when analyzing the data directly in the time domain. This is why in recent years, studies have been published (although few) that propose methods into how this may be accomplished. In an article published in 2001, Vanhamme et al [20] reviewed some of these current methods that employ model-fitting algorithms. Similar to model-fitting in the frequency domain, maximum likelihood estimates are made by minimizing the squared difference between the data and the model function, a typical nonlinear least-squares problem. However, the downside to these theories is that they are computationally inefficient. For this reason, good starting values are estimated using a process called 'peak picking'. These peak picking methods are termed as 'interactive', and the first most widely used was VAPRO, *va*riable *pro*jection method (for further details on VAPRO, consult reference [21]). The second of these, an improvement of VAPRO, is AMARES (*a*dvanced *m*ethod for *a*ccurate, *r*obust, and *e*fficient *s*pectral fitting) [22]. Vanhamme also describes 'black-box' methods which are

beyond the scope of this thesis, and interested readers can refer to the review publication [20].

Another time-domain analysis method uses adaptations of the Generalized Rank Annihilation Method (GRAM) [23-25] and the Direct Exponential Curve Resolution Algorithm (DECRA) [3, 26, 27], both model-free schemes. The latter of these two methods is described in detail in chapter V. Both methods convert the FID signal into a Hankel matrix and retrieve quantitative data from then on. However, the GRAM method compares the analyte Hankel matrix to that of a reference mixture. DECRA uses a single sample FID and separates it into two matrices, resolving individual component concentrations. Nordon et al [3, 27] compared these two methods and showed that DECRA had significant advantages over GRAM for quantitative analysis by low-field NMR spectroscopy. The main advantage was obtaining accurate (< 5% error) and precise (2.3% relative standard deviation) quantitative results.

2.4 Advantages to Quantitative NMR Spectroscopy

With so many parameters that require optimization, why is NMR spectroscopy a continuous focus of quantitative research, if other methods exist that do not involve as many experimental preparations? The advantages of using NMR spectroscopy as a quantitative analytical tool are in fact numerous. Because of the simultaneous qualitative and quantitative information gained from a single NMR signal, NMR spectroscopy proves to be an invaluable analytical tool. An important aspect of NMR spectroscopy is that the fundamental physicochemical mechanism is completely different from other common analytical techniques and serves to provide a different scientific perspective [2].

In particular, NMR spectroscopy is distinctly different from chromatographic procedures, which is one of the common current methods of choice for quantification. Chromatography employs the mantra of "separation means knowledge" and is limited to the chemical dynamic range [2], whereas NMR has the advantage of being a non-destructive tool, and is limited only by magnetic field strength.

NMR also differs from most spectroscopic quantitative methods, such as Raman or infrared spectroscopy, as they normally require external calibrations [28]. NMR has the potential to determine relative quantities of resolved pure components without the need for these and other calibration techniques, unless absolute concentration determinations are desired. Moreover, because of large spectral width (*sw*) possibilities (i.e. the range of nuclear precession frequencies to be examined [4]), multiple analytes can be analyzed simultaneously. Whereas common biological fluid quantitative analyses require individual analyte biochemical assays, especially for metabolites examined in amniotic fluid. This process can become very expensive when considering the numerous metabolites present.

This type of metabolic profiling of amniotic fluid can also be used as a means of continuous diagnosis of abnormalities *in utero* during pregnancy. The application of NMR spectroscopy as a non-invasive instrument has been accepted within the medical community with systems such as Magnetic Resonance Imaging (MRI) and Magnetic Resonance Spectroscopy (MRS) to monitor and diagnose abnormalities within the entire human body [29-33]. MRI provides anatomical information based on signals from water, and MRS serves to offer metabolic chemical information. Ideally, NMR could be used as a non-invasive monitoring tool, either spectroscopically or imaging-wise, implemented

continuously throughout gestation, particularly in pregnancies at increased risk [34]. As opposed to amniocentesis, which is normally only implemented once during pregnancy.

NMR spectroscopy sensitivity is also a key advantage to its applications to quantitative analysis of biological fluids. Most metabolites present at concentrations above ~0.1mM are detectable with ^{1}H NMR and in many cases quantifiable [14]. The equilibrium magnetization generated by an external magnetic field (signal intensity) increases linearly with strength of the external magnetic field, which results in greater sensitivity. So, as higher spectrometer fields and stronger magnets are being introduced, the resolution is increased, the sensitivity is continually pushed lower, and the signal to noise ratio (S/N) pushed higher [9, 35].

Equipped with the background of previous quantitative analysis methods for NMR spectroscopy, and knowing the instrumental requirements for a quantifiable NMR signal, the following chapter describes the calibration of quantitative NMR spectroscopy of biological solutions.

III. Calibration of Quantitative NMR

For metabolic profiling and metabolite measurements, a stable and accurate quantitative analysis method is needed. In this chapter, a quantitative NMR analysis method is developed. The chapter begins with an experimental section describing in detail the preparation and acquisition of quantifiable NMR signals for calibration biological solutions. The description is followed by an examination of quantitative analytical methods in both the frequency-domain and the time-domain. Limiting factors and sources of error for each routine were determined and optimal values were found. Based on simulation studies, a calibration of the NMR was determined by applying these analytical approaches to solutions of known biological constituents. The most accurate method was defined as a model protocol for routine ^{1}H NMR quantitative analysis of biological fluids.

3.1 Calibration Biological Solutions

Solution Preparation

To mimic amniotic fluid, a target biological solution standard was developed with two constituents found in amniotic fluid at high concentrations. A calibration set of six solutions was prepared by varying concentrations of alanine and glucose. Both have resolved ^{1}H NMR signals. A stock solution was first prepared by dissolving 0.1855g of glucose and 0.1285g of alanine in 100.00mL of solvent. The solved was prepared by diluting 6.0mg of TSP (sodium 3-(trimethyl-2,2,3,3)-1-propionic acid-d_4) in 250mL of 10% D_2O/90% H_2O. Inclusion of 10% D_2O/90% H_2O introduced a deuterium lock to monitor magnetic field strength drift, without the loss of proton signals through

deuterium-proton exchange processes. Six solutions were then prepared by diluting appropriate stock solution volumes with the same solvent to a final volume of 25.00mL, resulting in concentration ranges of glucose between 0.82 – 9.48mM and alanine between 1.15 – 13.27mM. For each biological solution, three replicate NMR measurements were collected.

In addition to measurements of the prepared calibration biological solutions, NMR spectra were also acquired for the calibration biological solutions after a standard addition. The standard addition serves to determine absolute concentration using the classical integration method. A standard addition of 10μL 2.06mM of alanine (doublet centered at 1.48ppm in figure 3.1) was injected directly into each sample and a second set of NMR spectra was collected using the same parameters listed above.

Data Collection

Proton NMR spectra of each calibration biological solution (600μL) were measured on a 500MHz Varian INOVA NMR spectrometer. The water resonance at 4.8ppm was suppressed using excitation sculpting [36]. The delay time was set to 20s, to allow for sufficient relaxation of the nuclei to the original Boltzmann distribution prior to the 90° transient RF pulse. This RF pulse was set to 7.5μs; 4 FIDs were accumulated and summed; 16K data points were acquired, with a spectral width of 8000Hz. Data was processed using MestRe-C® software (version 441) and Matlab (version 6.5.1). Figure 3.1 is an example of the Fourier-transformed ^{1}H NMR spectrum of a calibration biological solution containing 13.27mM alanine and 9.48mM glucose. Resolved signals for both metabolites are visible in the spectrum, with the alanine doublet centered at

1.47ppm, and the glucose multiplet ranging from 3.16 – 3.96ppm. The remnant water signals and the chemical shift reference singlet of TSP at 0.00ppm are also visible.

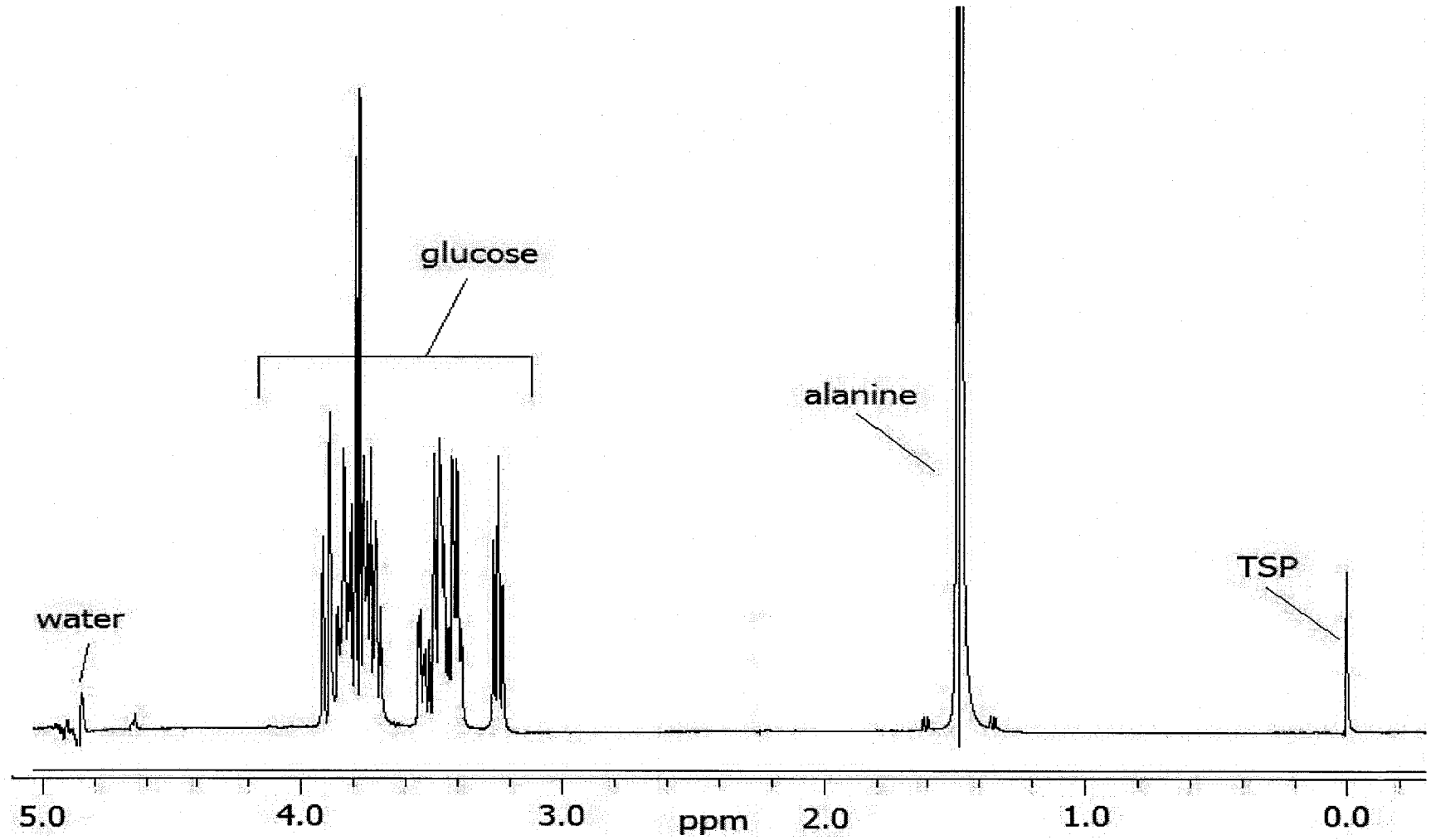

Figure 3.1 – Fourier transform ^{1}H NMR spectrum of a calibration biological solution of 9.48mM glucose and 13.27mM of alanine from a 500MHz spectrometer

3.2 Quantitative Analysis Calibration

3.2.1 Frequency-domain Quantitative Analysis

Integrated signal areas of constituents within the Fourier transformed NMR spectrum should be directly proportional to the quantity of the analyte examined within the sample. Signal integration coupled with a standard addition methodology has become a 'classical' quantitative analysis method in determining absolute concentrations. The theory behind the standard addition methodology is an extension of the fundamental relationship of Equation 2.1. Given that the NMR signal is proportional to the quantity of analyte within a given sample, a change in the NMR signal is a reflection of a change in the sample itself. In this case, the addition of analyte to the solution should be manifested as a linear increase in the specific analyte integral. For NMR, this approach was first applied by McGowan et al [13, 15], who obtained average standard error of 22.8% for AF metabolites quantified.

A single standard addition of alanine combined with the classical integration method was first used on the calibration biological solutions. The alanine signal was integrated between 1.42-1.56ppm (figure 3.1), before and after the addition of 10μL of 2.06mM of alanine (C_{SA}) to each 600μL biological solution, and the difference in signal integral was determined as $I_{difference}$. Glucose concentrations were also estimated using this standard addition quantitative method and the respective signal was integrated between 3.00-4.00ppm from the ^{1}H NMR spectra (figure 3.1). The concentrations of alanine and glucose for each triplicate measurement per biological solution, $C_{Sample,met}$, was estimated, compensating for the number of equivalent hydrogens associated with the metabolite's functional group (alanine $n = 3$; glucose $n = 2$).

$$C_{sample,met} = I_{sample,met} \times \frac{C_{SA}}{I_{difference}} \times \frac{n_{sample,met}}{n_{alanine}} \qquad 3.1$$

Triplicate alanine and glucose concentration estimates were then plotted versus their known concentrations. After a linear regression analysis of the data, an R^2 of 0.014 was obtained with a coefficient of variance (C.V.) of 281% for alanine, and an R^2 of 0.018 was obtained with a C.V. of 284% for glucose.

Standard addition analysis of the test solutions resulted in large quantification errors nearing 300%. The major source of the error is the inevitable variation in magnetic field strengths as a result of magnet shimming, which alters the phase and peak-shape of the NMR signal. Changes in the chemical environment as a result of standard addition dilution can also alter the target NMR signals. An example of these effects is shown in figure 3.2. The slight differences in magnetic field gradients over the sample, as well as the change in chemical environment, are shown by a slight change in chemical shift and alteration of peak-shape after the standard addition. Methods that could decrease the effects of phasing should provide better quantification accuracy of these NMR signals. Analysis of the time-domain signal FID avoids these associated errors by avoiding preprocessing techniques like chemical shift alignment and manual phasing.

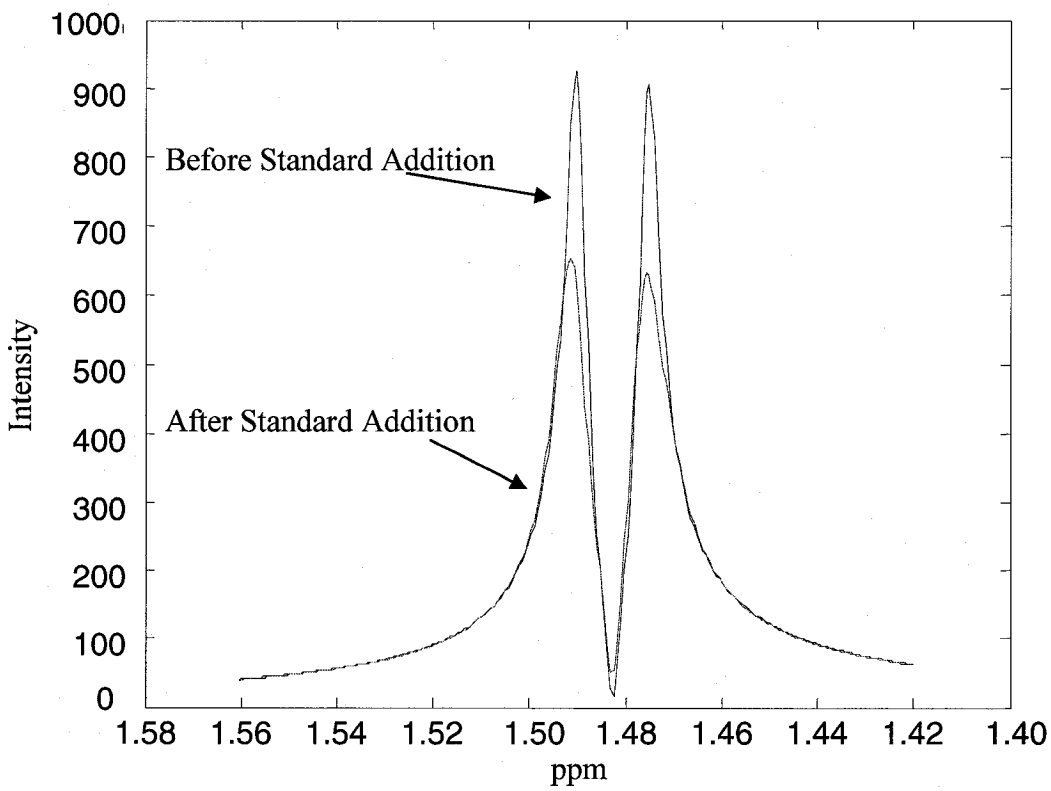

Figure 3.2 – Alanine NMR signal phase and chemical shift alterations as a result of variations in gradient shims and chemical environment before and after standard addition.

3.2.2 Time-domain Quantitative Analysis

Theory - Direct Exponential Curve Resolution Algorithm (DECRA)

Multivariate statistical approaches can be used on the time-domain NMR signals to avoid manual preprocessing phase and chemical shift error limitations. The direct exponential curve resolution algorithm (DECRA) method exploits the exponentially decaying nature of the FIDs to resolve the individual component contributions [26]. The DECRA approach is composed of a series of simple linear algebraic steps.

In general, an NMR data set (D) representing mixtures can be expressed as the superposition of individual pure spectra (P) weighted by their concentration (C), where

$$D = CP^T \tag{3.2}$$

The superscript T denotes the transpose of the pure spectra matrix. As FID as a function of time, $D(t)$, is a summation of K independent exponentially decaying sinusoidal waveforms, with the composition weightings given by amplitudes, A. Each sinusoidal waveform has a characteristic frequency of v_0, a phase angle of θ, and a transverse relaxation time of T_2. Using the FID relationship, Equation 3.2 can be alternatively expressed as

$$D(t) = \sum_{k=1}^{K} A_k \exp(i[2\pi v_{0k} t + \theta_k]) \exp\left(\frac{-t}{T_{2k}}\right) \tag{3.3}$$

The first exponential component of Equation 3.3 describes the pure spectral components associated with the kth component, and the second exponential component dictates the exponentially decaying character of the signal. Direct solution of Equation 3.2 is ill-conditioned, since there are an infinite number of solutions. However, due to the repetitive form of Equation 3.3, a unique solution can be found. Kubista and Scarminio [37] have shown that one solution can be obtained if the data can be separated into which two subsets are proportional, so that

$$A = CP^T \tag{3.4}$$

$$B = C\beta P^T \tag{3.5}$$

The matrix β is a diagonal matrix, the diagonal elements of which scale the pure component spectra and concentrations between A and B. Finding a unique β is the challenge for this approach. However, the matrix forms of Equations 3.4 and 3.5 lend themselves for solution by applying the method of rank annihilation (GRAM) [38-40].

To understand the GRAM approach, it is useful to first invert Equations 3.4 and 3.5, to obtain

$$C = A(P^T)^+ \tag{3.6}$$

$$C\beta = B(P^T)^+ \tag{3.7}$$

where $(P^T)^+$ represents the pseudo-inverse of the matrix P^T. Multiplying matrix β to both sides of Equation 3.6, results in

$$C\beta = A(P^T)^+\beta \tag{3.8}$$

Equating Equations 3.7 and 3.8, forms a generalized eigenproblem, given by

$$A(P^T)^+\beta = B(P^T)^+ \tag{3.9}$$

In the form of Equation 3.9, there is only one possible solution. This solution is found by noting that a common base between matrices A and B exists and can be obtained from a singular value decomposition (*svd*) of matrix B, where

$$svd(B) = USV^T \tag{3.10}$$

By substituting B in Equation 3.9 with the *svd* solutions in Equation 3.10, a standard eigenvector-eigenvalue problem is found, where

$$(P^T)^+\beta = (U^T AVS^{-1})(P^T)^+ \tag{3.11}$$

with $(P^T)^+$ defining the eigenvectors and β, the eigenvalues. The concentrations are then determined along with the respective pure spectral components by the matrix multiplication of

$$C = U(P^T)^+ \tag{3.12}$$

$$P^T = (VS^{-1}(P^T)^+)^+ \tag{3.13}$$

The composition weighting of the kth signal, A_k, (Equation 3.3) can be computed without interference from the other constituents in the mixture by taking the absolute value of the product of the first coefficients from the corresponding columns of C and rows of P^T [3],

$$A_k = \left| C_{1,k} \times (P^T)_{1,k} \right| \qquad 3.14$$

Implementing this DECRA method was initially done by Antalek and Windig [23, 26] using pulsed gradient spin echo (PGSE) NMR measurements [41]. PGSE data sets are comprised of multiple frequency-domain spectra collected as a result of varying the strength of two sets of magnetic field gradient pulses. However, Nordon et al [3] showed the same analysis can be applied to a single FID. The two data sets in Equations 3.4 and 3.5 can be constructed from a single FID because of its characteristic exponentially decaying profiles. Different sub-segments of an exponentially decaying profile are self-similar and only differ by the offset and scaling factor β.

To construct two data subsets A and B, a Hankel matrix, $\boldsymbol{H}$, of size N x N is first constructed from a single FID. A Hankel matrix is a square matrix with constant skew diagonals. In constructing a Hankel matrix from an FID of length N, each successive row is produced by dropping the first point from the prior row and adding a zero to the end of the row,

$$H = \begin{bmatrix} H_1 & H_2 & H_3 & \dots & H_N \\ H_2 & H_3 & H_4 & \dots & 0 \\ H_3 & H_4 & H_5 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_N & 0 & \dots & \dots & 0 \end{bmatrix} \qquad 3.15$$

$\boldsymbol{H}$ is partitioned such that matrix A is produced by dropping the first I rows of $\boldsymbol{H}$ (I=1 normally), and matrix B is produced by dropping the last I rows of H. Therefore, A consists of rows 1:N-1 and B, 2:N. An example of the Matlab script used in this study is provided in [3]. The critical elements of the Matlab script used in this thesis are given in Appendix A.

The number of unique components that can be resolved using DECRA is dependent on the noise present. Typically, the number of unique components has been limited to four or less [42]. Initial application of DECRA was limited to few analytes from a single FID using low-field NMR spectroscopy [3]. Most biological fluids are complex mixtures, thus requiring analysis by a modified version of the DECRA method. One such approach to DECRA is to reduce the number of components analyzed by isolating them in the frequency-domain using a windowing function.

For the windowing approach, an FID (Equation 3.3) is Fourier transformed to a frequency-domain signal. The frequency-domain signal, $F(\varpi)$, is then multiplied by a windowing function, $W(\varpi)$. This windowing function is an array of 0's and 1's, of equal length to $F(\varpi)$. The windowing method can be achieved by a simple element-by-element multiplication defined as

$$Z(\varpi) = S(\varpi)W(\varpi) \qquad 3.16$$

The resulting signal, $Z(\varpi)$, is inverse Fourier transformed to result in a reduced complexity FID for DECRA analysis. This process can be visualized by referring to figure 3.3. In figure 3.3(a) a complex FID for the full spectral range is shown. Fourier transformation of this FID results in the frequency-domain spectrum given in figure 3.3(b). The window function isolates the region in the frequency-domain spectrum to be

analyzed by the DECRA method. The arrow of this window in figure 3.3(b) indicates the user defined changing width and position of the window. Window width and position are chosen depending on the chemical shift region of the constituent desired. This region is then inverse Fourier transformed with the rest of the spectral elements set to zero. The less complex FID (figure 3.3(c)) is then analyzed using DECRA, resulting in estimated pure spectral components and respective concentrations.

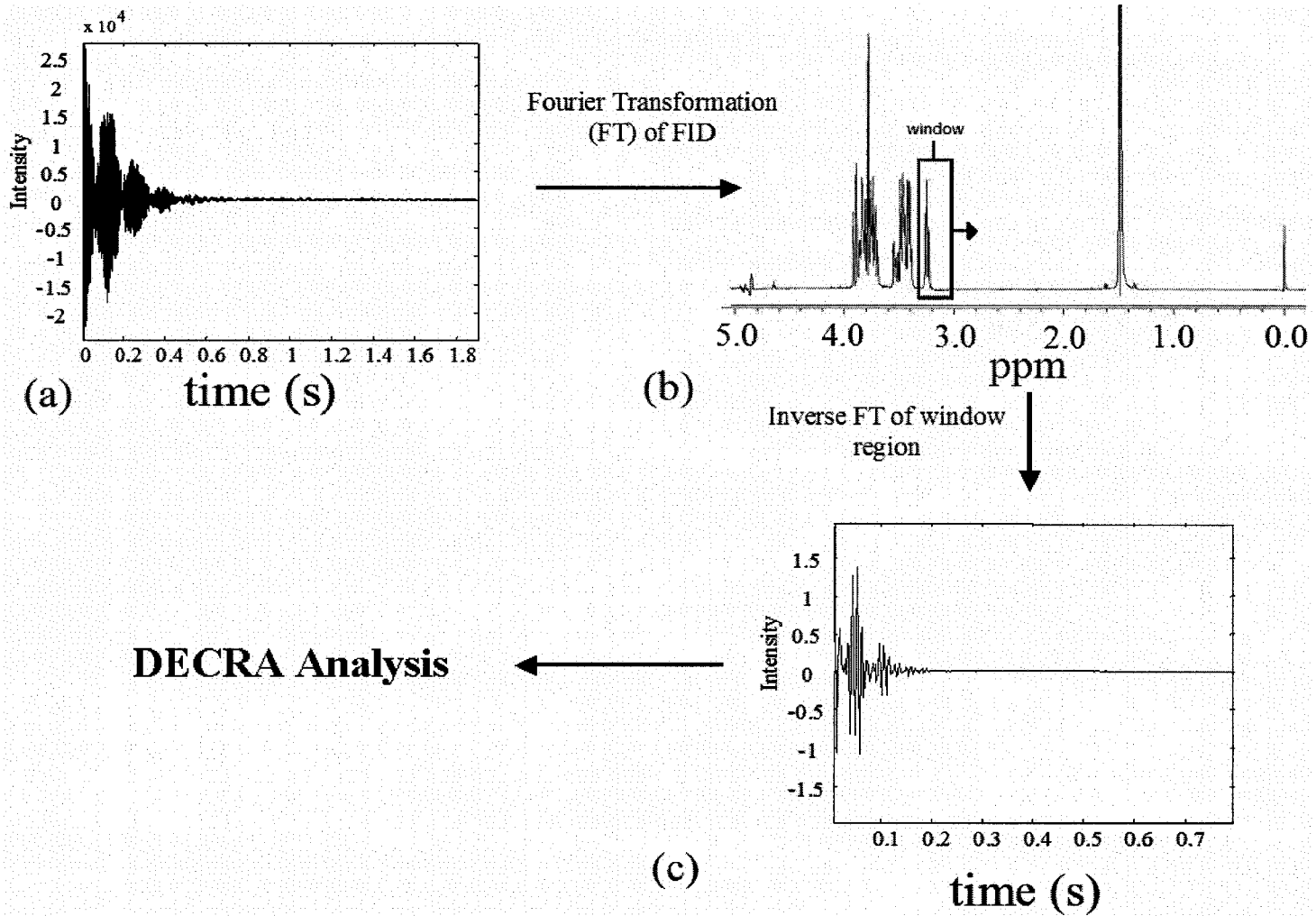

Figure 3.3 – Routine for windowed DECRA of Fourier transformed calibration biological solution ^{1}H NMR spectrum (a) The FID of the full spectrum is Fourier transformed (b) Specific chemical shift regions are windowed and inverse Fourier transformed (c) The resulting FID is run through the DECRA method resulting in pure spectra and concentration estimates of the windowed region.

Results – Direct Exponential Curve Resolution Algorithm (DECRA)

The DECRA method was tested on simulated and real data sets to determine the sensitivity and uncertainty of the method to common instrumental variations observed in NMR. First, simulated FIDs were used to test the sensitivity to three major NMR instrumental limitations. The DECRA method was then evaluated on the calibration biological solutions, resulting in a calibration quantitative NMR analysis method.

A. Simulated FIDs

Free inductive decays (FIDs) were simulated using Matlab to test the sensitivity of the DECRA method. From Equation 3.3, we can see the signal intensity as a function of time, $D(t)$, is correlated to numerous parameters including the signal amplitude, A, the transverse relaxation time, T_2, and the phase angle, θ. Changes in any one of these three parameters create distortions that are readily observed in the frequency-domain NMR spectrum. Therefore, to simplify visualization in this study, all effects are described in terms of frequency-domain distortions. Manipulation of these three parameters may affect the uncertainty associated with estimating concentrations using the DECRA method. Distortions tested include the effect of varying the line-width at half-height (T_2 relaxation time), baseline addition (amplitude, A), and baseline distortions (phase angle, θ). The predominant factor contributing to error between these three was determined.

1. Effect of Varying Line-Width

An improperly shimmed magnet can cause distortions in the line-shape of an NMR signal in the frequency-domain spectrum. Improper shimming affects the

transverse relaxation time (T_2) of the nuclei within the sample. Because each sample must be shimmed individually, different line-shape distortions appear in each sample NMR spectrum. This theory was tested on simulated FID data to determine whether an improperly shimmed magnet, and thus a skewed line-shape and variable T_2, would affect the concentration estimation of the DECRA method.

A set of FIDs were simulated (8192 points) containing summed sinusoidal waveforms with frequencies of $v_{01} = 755$ and $v_{02} = 765$ Hz, each having an amplitude of 100, phase angles of 0rad, and an addition of 10% random noise to the signal. The T_2 relaxation times for each frequency were varied between 0.05 – 1.10s, giving line-widths at half-heights between 20.0 – 0.9Hz, respectively. DECRA analysis was used on the resulting set of FIDs, where two components were resolved from each signal, along with their relative concentration estimates.

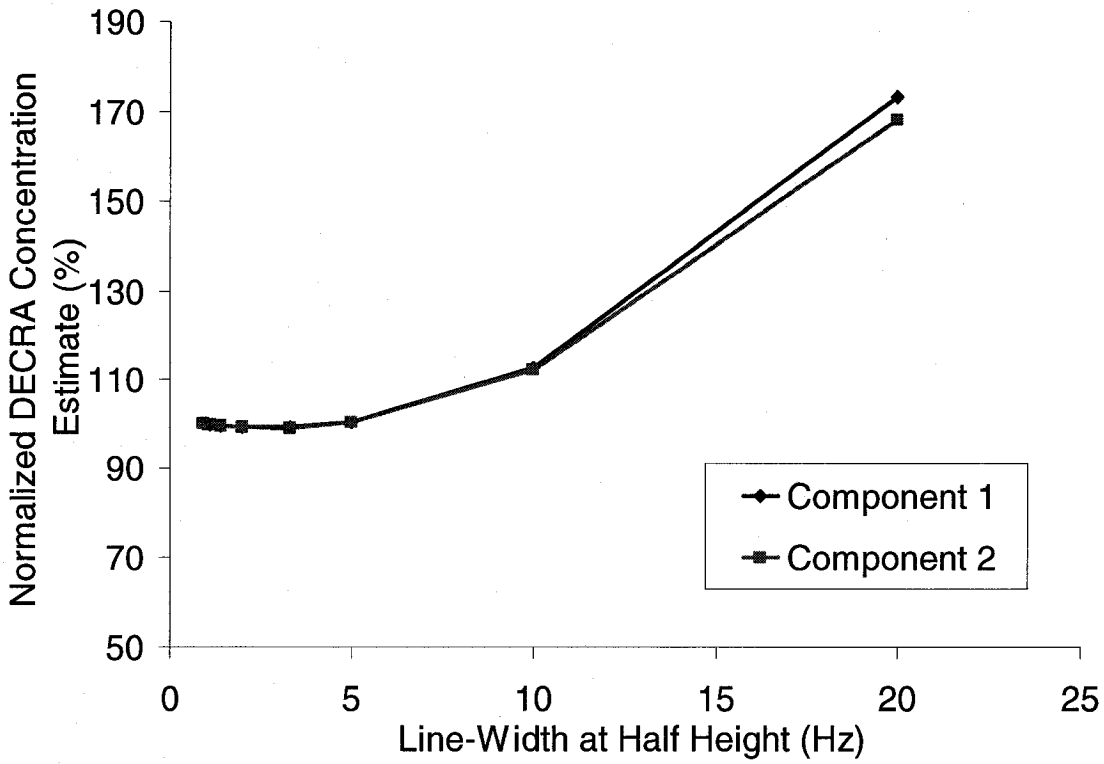

Figure 3.4 – Effect of varying line-width at half-height between 0.9-20.0Hz (T_2 relaxation time, 0.05 – 1.10s) on DECRA concentration estimate of simulated two-component FIDs

Figure 3.4 displays the sensitivity of the DECRA concentration estimates to line-width at half-height for each of the resolved signals in the simulated frequency-domain spectrum. As the line-width of a peak increased (i.e. as the T_2 relaxation time decreased), DECRA over-estimated the respective signal concentrations. From figure 3.4, it can be concluded that the accuracy of DECRA is dependent, in a non-linear correlation, on the T_2 relaxation time of an FID, and thus dependent on the shimming of a magnet prior to acquisition. However, little difference in the DECRA concentration estimate is observed if relatively small line-width differences exist. To reduce line-width variation, it is common to apodize the NMR signals by using a convolution function, and would be recommended for the DECRA analysis.

2. Effect of Variable Baseline Addition

To test the sensitivity of the DECRA method to non-zero baselines, FIDs were again simulated, containing two components with frequencies of $v_{01} = 755$ and $v_{02} = 765$ Hz, having amplitudes of 100, phase angles of 0rad, T_2 relaxation times of 0.5s and an addition of 10% random noise. These FIDs were Fourier transformed into the frequency-domain, where a Gaussian shaped signal, with a peak intensity of varying component amplitude percentages (0 – 10%), was added to the spectrum. This signal addition created a non-zero, non-uniform frequency-domain baseline. The total frequency-domain signal was inverse Fourier transformed into a time-domain signal and analyzed by DECRA.

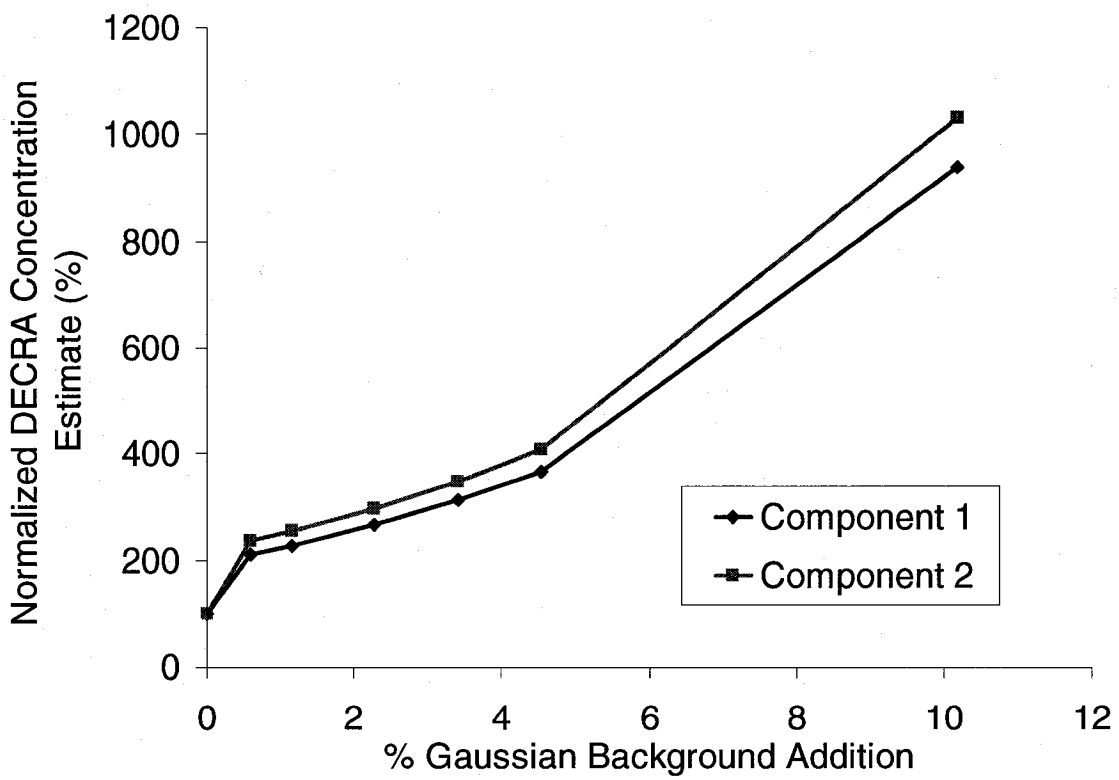

Figure 3.5 – Effect of % Gaussian baseline addition (0 – 10%) on DECRA concentration estimate of simulated two-component FID

Figure 3.5 shows the sensitivity of the DECRA concentration estimates to the percent Gaussian shaped baseline added. It is clear that the accuracy of DECRA is dependent on the baseline addition to the signal. In a similar sensitivity to line-width at half-height, DECRA overestimates the component relative quantities when a non-zero baseline exists. For accurate concentration estimations by DECFRA, a clearly resolved, zeroed baseline signal is required.

3. Effect of Varying Phase Angle

Manual and/or automatic zero-order and first-order phase adjustments are not possible with common NMR processing software for a time-domain signal, requiring the DECRA method to be insensitive to phase angle variances that may result from pre-acquisition and detection of the signal. To study the effect of phase angle on the accuracy of DECRA, the following parameters were used to simulate FIDs: two components with frequencies of $v_{01} = 755$ and $v_{02} = 765$ Hz, having amplitudes of 100, T_2 relaxation times of 0.5s and an addition of 10% random noise. The phase angles were sequentially changed with values of 0, $\pi/2$, π, and $3\pi/2$ rad for each frequency component.

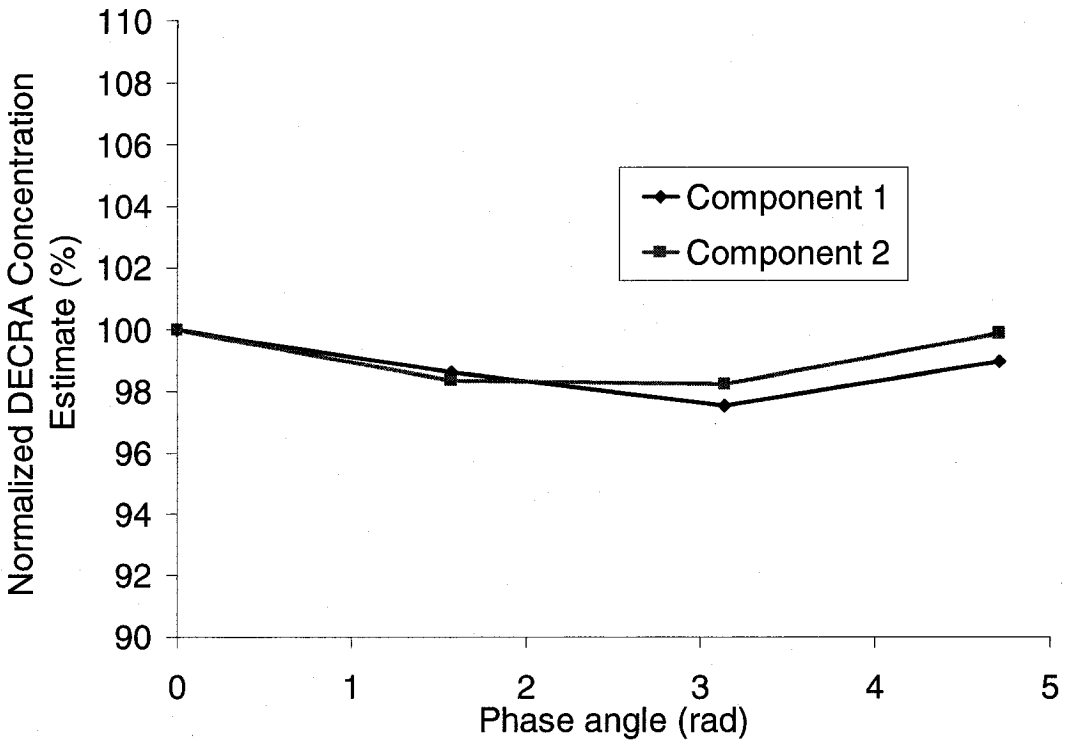

Figure 3.6 – Effect of phase angle (0, $\pi/2$, π, $3\pi/2$ rad) on DECRA concentration estimate of simulated two-component FID

The simulated FIDs were analyzed by DECRA, and figure 3.6 shows little sensitivity of the DECRA concentrations estimated to phase angle variations of a time-domain signal. Relative concentration estimates remained the same, with approximately 1% error between varying phase angle values. Results suggest that the accuracy of DECRA is independent on the phase angle of the time-domain NMR signal and neither zero-order nor first-order phase corrections need be applied to the data prior to analysis with DECRA.

B. Calibration Biological Solutions

Knowing DECRA's sensitivities to common NMR signal limitations, the method can now be applied to FIDs collected for calibration biological solutions. Applying the

windowing DECRA analysis routine of figure 3.3 to these solutions will determine the linearity and accuracy of the analysis method. The frequency-domain spectra of all calibration biological solutions were first obtained by Fourier transformation of their respective FIDs. Both alanine and glucose regions were analyzed separately by isolating each metabolite between 1.15 – 1.75ppm for alanine, and 3.05 – 3.35ppm for glucose, using the windowing method. Regions of the frequency-domain spectra that are strictly baseline signal were isolated (1.10 – 1.17ppm for alanine, 2.35ppm – 2.50ppm for glucose), and an average baseline value was computed for each spectrum and subtracted from the total signal. Spectra were convoluted with a Lorentzian function to limit any change in line-width. In the frequency-domain the convolution function was given as

$$g(v) = \frac{1}{\pi}\left(\frac{\sigma}{(v-v_0)^2+\sigma^2}\right) \qquad 3.17$$

where σ represents the half line-width at half-height of the Lorentzian curve (similar to a standard deviation in a normal distribution), and v_0 is the peak maximum location. Spectra were apodized using the Lorentzian function (Equation 3.17), where σ was given by 3 times the largest alanine peak line-width from the doublet centered at 1.47ppm. The resulting spectra were inverse Fourier transformed and analyzed using DECRA.

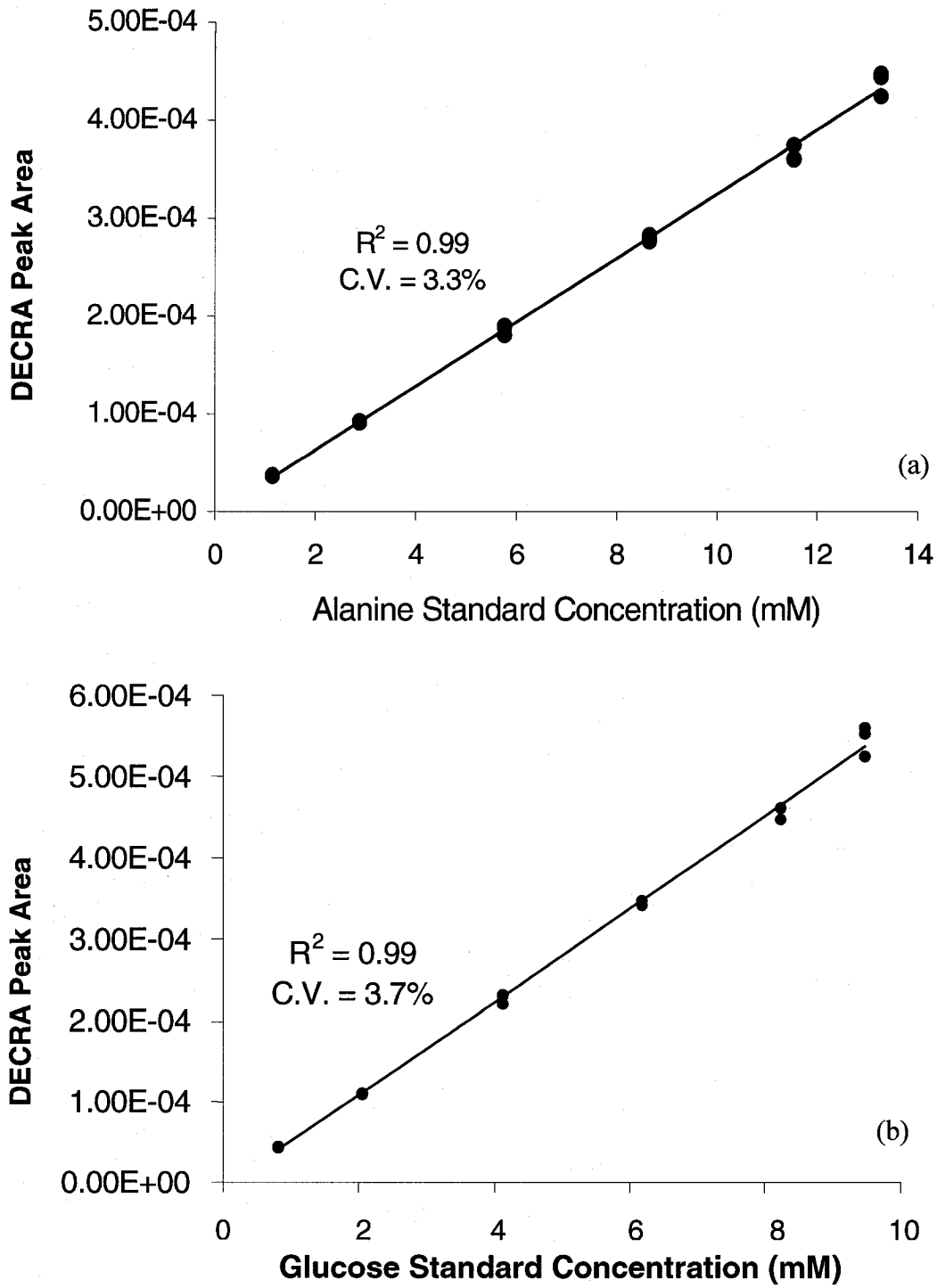

Figure 3.7 – (a) Triplicate DECRA peak area estimate for alanine concentrations isolated between 1.15 – 1.75ppm, after apodization with a Lorentzian shape curve (3×largest alanine peak line-width) R^2=0.99 and C.V = 3.3% (b) Triplicate DECRA peak area estimate for glucose concentrations isolated between 3.05 – 3.35ppm, after apodization with a Lorentzian shape curve (3×largest alanine peak line-width) R^2=0.99 and C.V. = 3.7%

Peak areas of the alanine and glucose regions isolated, using the window function and estimated by the DECRA method, were plotted versus their known concentrations in figures 3.7(a) and 3.7(b), respectively. A coefficient of variance (C.V.) improved five times for alanine and 2 times for glucose after correcting the factors that limited the DECRA sensitivity. This calibration, using prepared biological solutions, has demonstrated that a robust method for quantitative NMR analysis now exists, and can be applied to real biological fluids.

3.3 Quantitative NMR Protocol Summarized

The ultimate goal of the research design was to determine a protocol for ^{1}H NMR quantitative analysis of biological fluids. To do this, factors that limited the accuracy of this measure were determined and corrected for using a calibration method. Between three NMR signal distortions, the line-width at half-height (T_2 relaxation time), baseline addition (amplitude, A), and phase angle (θ), the predominant factor affecting the DECRA accuracy can be concluded to be the line-width of an NMR signal in the frequency-domain. When analyzing any NMR data set, this factor must be corrected prior to DECRA analysis. The concise windowing DECRA protocol is described below, beginning with the collection of the NMR signal to the resulting concentration estimates.

Quantitative ^1H NMR windowed DECRA protocol

1. A quantifiable ^1H NMR time-domain signal is collected from a high-resolution NMR spectrometer. Adjusting optimal acquisition parameters ensures a signal that obeys the fundamental relationship outlined by Equation 2.1.
2. The time-domain signal is Fourier transformed to the frequency-domain ^1H NMR spectrum.
3. A window function is multiplied with the frequency-domain spectrum, isolating key metabolite frequency regions.
4. The line-widths of largest peaks within the isolated metabolite frequency regions are determined.
5. A Lorentzian curve of 3 times the line-width determined in step 4 is convoluted with the respective isolated metabolite frequency region, according to Equation 3.16.
6. The regions are inverse Fourier transformed resulting in time-domain FIDs.
7. These FIDs are analyzed by the DECRA method, resulting in concentration estimates of the key metabolites isolated.

IV. Amniotic Fluid and Physical Health

Many biological fluids can produce metabolic signatures of biological systems, and can be used to monitor the metabolic processes of the specific organism. Amniotic fluid (AF) provides an interesting example of a biological fluid that can result in metabolic signatures, which represents the physiological processes occurring throughout pregnancy. As AF has contributions from the mother, the fetus, and the placenta, metabolic profiling is used to monitor the clinical health of both the mother and the fetus. In this chapter, the background literature of the nature of AF and methods to acquire it are discussed. Likewise, the thesis focuses on a target metabolic disorder that occurs during pregnancy, gestational diabetes mellitus (GDM). Literature pertaining to the prevalence and current methods of diagnosing GDM is also given.

4.1 Amniotic Fluid

Amniotic fluid (AF) is the liquid that surrounds a fetus within the amniotic sac during gestation. The fluid serves to cushion the fetus and maintain a stable temperature and pressure throughout the pregnancy term. The origin of AF has been presented by many theories and experimental evidence. However, over the past 50 years, it has been shown that AF is not a stagnant pool but is in constant circulation [43, 44] and is continually renewing and exchanging across surrounding membranes. The origin of AF seems to vary at different points throughout gestation, yet the fluid is ultimately derived from both the mother and the fetus. Earlier on, AF results primarily from maternal secretion from the placenta and other membranes. Whereas further along during gestation, AF is derived from the fetal gastrointestinal tract, skin, kidneys, tracheobronchial tract and other fetal

organs as they mature and are able to absorb and secrete the fluid [45]. Because of this, AF constituents at different points throughout pregnancy are reflective of the clinical state of the fetus and/or mother and can be used to diagnose pathological and physical conditions experienced during (prenatal) and post-pregnancy.

4.2 Methods of Prenatal Diagnosis

Diagnosing maternal and fetal health prenatally has evolved from simple ultrasonographic methods to complex screening routines. Today, patient health is assessed primarily through chromosomal abnormality analysis (karyotyping) and metabolic profiling (metabonomics and metabolomics) of AF. Both of these methods reveal different perspectives concerning maternal and fetal clinical statuses. Chromosomal analysis can show abnormalities in the chromosome as a result of genetic disease or malformation. Metabolic profiling, defined as the qualitative and/or quantitative measure of metabolites in a biological system, can reveal changes due to pathological and/or physical stimuli. All known chromosomal irregularities, including 120 recessively inherited diseases can be identified by measuring the activity of a specific enzyme or metabolic abnormalities resulting from its deficiency [46]. For example, lipid (lecithin to sphingomyelin, L/S ratio) measurements serve to identify lung maturity [46-49]. Birth weight classes can be estimated by monitoring the concentrations of glucose, insulin and specific insulin-like growth proteins [50-57]. Levels of AF glucose, alanine, lactate and insulin serve as biomarkers for the occurrence of gestational diabetes mellitus (GDM) [53, 58-61].

Theoretically, these screening methods are simple. However, implementing these tests is not trivial. Each test entails multiple steps and results may be difficult to interpret.

Waiting periods for expecting parents can be as long as two weeks and can cause unwanted stress on the mother, which can physically affect both the fetus and the mother. Variability and uncertainty in these test results can also provoke unnecessary treatments, which may also be hazardous. Many tests routinely implemented require a sample of AF. Though there are several methods to obtain amniotic fluid, the most common method involves amniocentesis.

4.2.1 Amniocentesis

Both karyotyping and metabolic profiling requires a sample of AF, which is obtained first by amniocentesis. This technique is where most of the risks lie when screening for gestational diseases. Amniocentesis requires a highly skilled and experienced technician and implicates both maternal and fetal risk. Normally, the procedure is done during the second trimester, between 15-20 weeks gestation and is generally carried out via the abdominal route; however, the vaginal route is also in practice. This discussion is limited to the first of the two methods because it is the preferred route after the 14^{th} week of gestation [46].

Procedure

Not every pregnancy undergoes amniocentesis. The physician assesses the need for the test by noting maternal age, medical history, and previous pregnancy complications. The procedure is then initiated with a skilled ultrasonographer locating the placenta and the position of the fetus. An ideal site is chosen, normally between fetal arms and legs (if developed) or the nape of the fetal neck. A 19/21 gauge spinal needle is then inserted through the abdominal and uterine walls into the amniotic sac (average

depth of 3-4 cm) and the AF is extracted by use of a syringe. Because multiple needle insertions increase the risk of hazards, the technician must insert the needle at a proper angle and depth on the first attempt. Ten to thirty milliliters are extracted and the color and the presence of blood-staining are noted. The sample is then sent to appropriate laboratories for testing [45, 46].

Risks and Complications

As most medical procedures, second trimester amniocentesis has a number associated of risks. There exist possibilities of maternal hemorrhaging, infection, and increase of maternal sensitization [45]. A European study between 1994 and 1997 with 8389 pregnancies (3091 cases and 5298 controls) [62], showed that women having amniocentesis delivered premature infants (born between 22 to 36 weeks gestation) at an occurrence of 1.59%. However, the study went on to say that "mother's reproductive history and medical factors play a role in the occurrence of spontaneous or induced pre-term deliveries". Yet there still remains confirmation between amniocentesis and pre-term births in the two aforementioned groups.

Another major issue linked to amniocentesis is fetal loss. A study including 6457 patients showed "an aggregate rate of pregnancy loss of 1.4% after 16 weeks gestation but before 28 weeks" [63]. Other fetal injuries include respiratory distress and muscoskeletal malformations [64]. Fetal damage caused by needle puncture has occurred but remains rare. Most normal fetuses are kept out of the way of the needle, but in the case of puncture, fetal hemorrhaging, loss of an eye, and scars in the skin of an infant at birth have been reported [45]. With these risk factors in mind, it has nonetheless been stated that "mid-trimester amniocentesis remains the safest invasive procedure" [65]. Still

because of the risks and complications associated with amniocentesis, there exists a vigorous hunt for a non-invasive replacement for this technique.

4.2.2 Diagnosis of Gestational Diabetes Mellitus

One gestational pathological conditions of great importance is gestational diabetes mellitus (GDM). GDM is a maternal glucose intolerance condition first detected during pregnancy [58]. This definition is applicable regardless whether insulin is used to treat the condition, if diabetes pre-existed pregnancy, or if diabetes persists after the pregnancy [66]. After delivery, 90% of women with GDM return to their normal glucose tolerance level. Among women with previous history of GDM, 60% will develop it again in subsequent pregnancies. In addition, 40% of women with GDM will develop type 2 diabetes within 15-17 years after delivery [66].

Although the cause of GDM remains relatively unknown, it is thought that hormones produced during pregnancy (human placental lactogen (hPL), free cortisol, progesterone and prolactin) inhibit insulin activity, leading to the inability of the human body to process glucose [59]. High levels of glucose cross the placental membrane, whereas insulin crossing is inhibited. This leads to fetal hyperglycemia, stimulating the fetal pancreas, resulting in fetal hyperinsulinemia as well [67].

GDM occurs in approximately 5% of the population, making it one of the most common complications of pregnancy [66]. Those at risk for developing GDM are subdivided into high, medium and low risk classes [58]. Those at high risk have marked obesity, previous infants with macrosomia (abnormally large newborns, >4500g), have first-degree relatives with diabetes or a history of glucose intolerance. Low risk patients are normally less than 25 years of age, belong to a low-risk ethnic group (Hispanic,

Black, Native American, South or East Asian, Pacific Islander, or Indigenous Australian), do not have a first-degree relative with diabetes, experience normal pregnancy weight and weight gain, and do not have a history of abnormal blood glucose concentrations. Average risk patients lie in between the high and low risk categories.

Risks and Complications

The onset of GDM comes with an increased clinical risk. Major birth defects are possible, however this only appears in infants with mothers with severe hyperglycemia (AF glucose levels >6.7 mM) [58]. Stillbirth, where an infant is dead upon birth, has historically been an important complication of diabetic pregnancies. Recent large clinical studies [58, 67] have shown that careful monitoring of fetal movement and fetal heartbeat (cardiotocography, CTG) throughout gestation and especially between 32-34 weeks, lead to no excess of perinatal mortality. This type of monitoring allows a physician to properly treat a patient by adjusting their diet, or by insulin injections, reducing the risk of stillbirth.

Other risks include macrosomia, jaundice, respiratory distress syndrome, polycythemia (increase in the red blood cell mass), and hypocalcemia [51, 58, 67]. Cesarean deliveries among women with GDM are near double those for non-diabetic women, with 30% of GDM births are by cesarean section, as compared to the average rate of 18-20% [66]. This increase may be due to an increased rate in macrosomic infants.

Gestational Diabetes Mellitus Screening

The most common method of screening for GDM is a process in which the low-risk and high-risk patients are first distinguished and divided, and appropriate monitoring and tests are issued for each patient group [58].

The first GDM evaluation applied to the high-risk group, the oral glucose challenge test, is an initial assessment made at the first visit to the clinician and has a sensitivity of 90% [58]. The patient is asked to quickly drink a liquid containing 50-g of glucose. A blood sample is collected one hour after the liquid has been ingested. Normally, blood glucose levels peak within an hour and then begin to drop [68-72]. Blood glucose measures of 7.2mM and above indicate the presence of GDM, and the patient is recommended to have the complete glucose tolerance diagnostic test [58].

This complete glucose tolerance diagnostic test is similar to the oral glucose challenge test in that a given amount of glucose is consumed, and the glucose levels within the blood are monitored throughout a period of time. The patient is asked to fast from food overnight, prior to the day of testing. A blood sample is first collected when the patient arrives, and provides a baseline for comparing the subsequent blood glucose concentration values. Again, the patient is asked to quickly drink a liquid containing 100g of glucose. Blood samples are then collected at 1, 2, and 3-hour intervals after ingestion [68-72]. The following table 4.1 lists the limits of blood glucose measured at these time intervals that if were equal to or greater than, would conclude in the diagnosis of GDM [58].

Table 4.1 - Diagnosis of gestational diabetes mellitus after the complete glucose tolerance diagnostic test. Greater or equal blood glucose levels than those listed in the table are diagnosed with GDM

Time of Measurement	Glucose Concentrations (mM)
After overnight fast	5.27
After one hour	9.99
After two hours	8.60
After three hours	7.77

If the patient is diagnosed with GDM at this point, and are at high-risk for overt hyperglycemia, they are advised to undergo immediate AF glucose tests through amniocentesis. Low-risk patients are monitored through their gestational period and are later reassessed between 24-28 weeks gestation [58].

Amniotic Fluid Metabolites and Gestational Diabetes Mellitus

AF metabolic profiling of pregnancies diagnosed with GDM has been the topic of many reports [15, 53, 58-61]. Glucose has been identified as the primary biomarker for determining the onset of GDM. As seen in the previous sections, maternal blood glucose concentrations are indicative of the presence of GDM in a pregnancy. Yet, how AF glucose concentrations correlate to GDM remains a topic of disagreement. In 1982, Iyengar et al showed that AF glucose values decrease progressively with advancing gestation, but in diabetic women, significantly high levels of AF glucose were found as compared to patients with normal pregnancies [73]. Shiffman et al agreed with this when they reported that elevation of glucose concentration in AF after second trimester amniocentesis is predictive of subsequent abnormal glucose challenge test, and possibly GDM [74]. In a similar study conducted by Weiss et al [75], AF glucose increased by a

total of 42% in patients with abnormal glucose tolerance as compared to normal pregnancies. Weiss et al also concluded that insulin dependent diabetes mellitus (IDDM) had glucose values elevated by a total of 77%. In a contradicting report that same year, Kalhan et al [60] showed that neither glucose nor alanine metabolisms varied between pregnancies with GDM and normal pregnancies during fasting. However, the study was not repeated during the fed state, but was predicted to be related to pregnancy outcomes of patients with GDM. This study, as well as that done by McGowan et al [15], showed that both glucose and alanine change collinearly. McGowan's investigation also reported that AF alanine, glucose, acetate, and lactate were all found to have lower concentrations in IDDM pregnancies as compared to normal pregnancies. However, this study did not include pregnancies with GDM.

In later years, reports studied levels of insulin in AF as predictive markers for GDM. Carpenter et al [53] showed that elevated second-trimester insulin in AF is related to subsequent diagnosis of GDM. However, this study did not look at glucose as a co-variable link to GDM. A study conducted by Star et al [61] asked this very question: which analyte best predicts future maternal glucose intolerance and subsequent GDM, AF glucose or insulin? Star concluded that elevated AF insulin is in fact correlated with a rise in AF glucose. However, insulin appeared to be a better predictor of subsequent maternal glucose intolerance and GDM than glucose.

To resolve the aforementioned debating reports, and to determine which of the above metabolites correlates best with gestational diabetes, the following study was designed. AF specimens sampled at earlier points of gestation were analyzed and quantified using NMR spectroscopy and the protocol outlined in chapter III.

V. Application of Quantitative ^{1}H NMR to Amniotic Fluid

Amniotic fluid (AF) provides an ideal biological matrix to analyze with the quantitative ^{1}H NMR protocol developed in chapter IV. There are three contributors to AF at different points throughout pregnancy: the placenta, the fetus, and the mother. The goal of this chapter is to apply the developed protocol to examine a gestational condition affecting the mother, which is reflected in amniotic fluid in late pregnancy. Quantitative measurements of AF constituents will be examined at an earlier phase of gestation, approximately 16 weeks. This metabonomic study aims to determine if significant differences in AF constituent concentrations are present, at earlier stages of pregnancy, for normal patients compared to those who develop gestational diabetes mellitus (GDM).

The chapter begins with a description of patient recruitment, AF sample collection and exclusion criteria used to develop an appropriate data set for analysis. The examination continues with the methods outlined for NMR spectral acquisition, and for quantitative and statistical analysis. Results are then presented and discussed. This work was done in collaboration with Dr. Kristine Koski of the School of Human Nutrition and Dietetics, at McGill University, Montreal Canada.

5.1 Amniotic Fluid Collection and Selection

Patient Recruitment

AF specimens used in this study were obtained from routine clinical analyses via second trimester amniocentesis from patients at St Mary's Hospital center. Three hundred subjects were recruited from the hospital between 2000 and 2003, and were asked, via second signed consent, to give away the leftover or discarded portion of their

amniotic fluid. A recruiter explained the study objectives to the participants face to face, replied to their questions, and clarified their concerns in four different languages (English, French, Arabic and Spanish). If the patient agreed to participate, they were asked to fill out a questionnaire and sign a consent form. The questionnaires were written in both English and French. When necessary, translation was provided by the recruiter to Spanish or Arabic. The accompanier of the pregnant woman provided additional language translations. The questionnaires were filled out either by the patient, by the recruiter, or by the patient's accompanier, but the patient themselves signed the consent form. The AF samples to be tested were taken to the Montreal Children's Hospital. After completion of genetic testing, leftover portions of the normal AF samples were sent to McGill University for research purposes. Note that the AF specimens were then received, as part of a larger ongoing study, "The Role of Amniotic Fluid Constituents in Predicting Human Fetal Growth", from the lab of Dr. Kristine G. Koski at the School of Dietetics and Human Nutrition, McGill University, Montreal, Canada.

Ethical Considerations

The study was approved by the McGill Institutional Reviews Board through the Faculty of Medicine, and by St. Mary's Hospital center. Ethics approval forms granted by McGill University are included in Appendix B.

Subject Confidentiality

Confidentiality was maintained by the numerical coding. Numbers were assigned to the subjects and they were used in place of names on all assessment materials. The list linking numbers with names and data was stored in locked and secure files in McGill

University, Macdonald campus. The information obtained was accessible only to the research staff, and no subject was identified in any report of the project. The staff contributing to the study was aware of the ethical issues associated with this research, with specific attention to confidentiality.

Potential Risks to Subjects

There are no known risks associated with participation in this study, since the physician did not extract additional fluid for the study purpose. The amount of the amniotic fluid used in our study was normally discarded after genetic testing.

Consent forms

Patients who underwent amniocentesis had to sign two consent forms. The first form was agreeing to undergo the amniocentesis with no third party participation. The second form was to give the author and co-researchers (the third party) their authorization to collect the discarded portion of the amniotic fluid having no genetic abnormalities. By signing the consent form, permission was given to retrieve the normally discarded portion of the fluid from Montreal Children's Hospital Centre (MCH).

Collection of Amniotic Fluid Samples

After undergoing the routine amniocentesis at St. Mary's Hospital Center, the amniotic fluid samples were taken to the MCH for genetic testing. Collection of amniotic fluid samples for which consent was obtained, was completed every month from the MCH. After being transported in a cooler of dry ice, the AF samples were directly put in

the -80 ° C freezer in the laboratories. The volume of individual samples ranged between 5 and 20 ml, depending on the volume taken at amniocentesis. Those samples were arranged by name of subject and amniocentesis date.

Inclusion and Exclusion Criteria

The participant had to have a singleton pregnancy, a normal gestation, and no fetal complications. Pregnant women having pregnancy or fetal complications, like abnormal fetuses with chromosome abnormalities, or multiple pregnancies were excluded from the study on the basis that these factors may interfere with the results of the study.

Gestational Diabetes Mellitus Diagnosis

Clinical information for each pregnancy was obtained including the diagnosis of gestational diabetes mellitus. Of the 258 AF samples analyzed, 35 had been diagnosed with GDM.

5.2 Materials and Methods

NMR Spectral Acquisition

Samples for ^{1}H NMR analysis were prepared by diluting 45μL of amniotic fluid with 600μL of solvent. The solvent was prepared by diluting 6.0mg of TSP (sodium 3-(trimethyl-2,2,3,3)-1-propionic acid-d_4) in 250mL of 10% D_2O/90% H_2O. The inclusion of 10% D_2O/90% H_2O introduced a deuterium lock to monitor magnetic field strength drift, without the loss of proton NMR signals through deuterium-proton exchange

processes. The dilution of amniotic fluid with the large solvent volume allowed appropriate sample volume adjustment for NMR analysis.

Free inductive decays (FIDS) of the AF samples were acquired on a 500 MHz Varian INOVA NMR spectrometer with a HCN triple resonance cold probe having Z-axis gradients. The experimental temperature was maintained at 20.0 ± 0.1°C. Inversion recovery experiments (refer to section 3.2, 'Relaxation Delay') were implemented on test AF samples in order to determine the longest longitudinal relaxation time (T_1(max)) for the multiple proton signals in the spectrum. For each sample, a delay time of 20 seconds was set to allow all proton nuclei to be sufficiently relaxed ($5\times T_1$(max)) before the subsequent 90° pulse, whose length was set at 7.5μs. To enhance low concentration metabolite signals, suppression of the intense water peak at approximately 4.8ppm was accomplished using a pulse sequence outlined by Hwang et al [36], termed 'excitation sculpting'. For each spectrum, 4 scans were averaged and 16K data points were acquired in a spectral width of 8000Hz, with a total acquisition time of 3.5 minutes per sample. The FIDs collected were Fourier transformed and key metabolite signals were assigned by comparison with literature high-resolution NMR of AF, and is summarized in table 5.1 [76, 77]. A sample AF spectrum resulting from the above experiment is shown in Figure 5.1, where well-defined metabolites are visible in the spectrum between 0.00 – 5.00ppm.

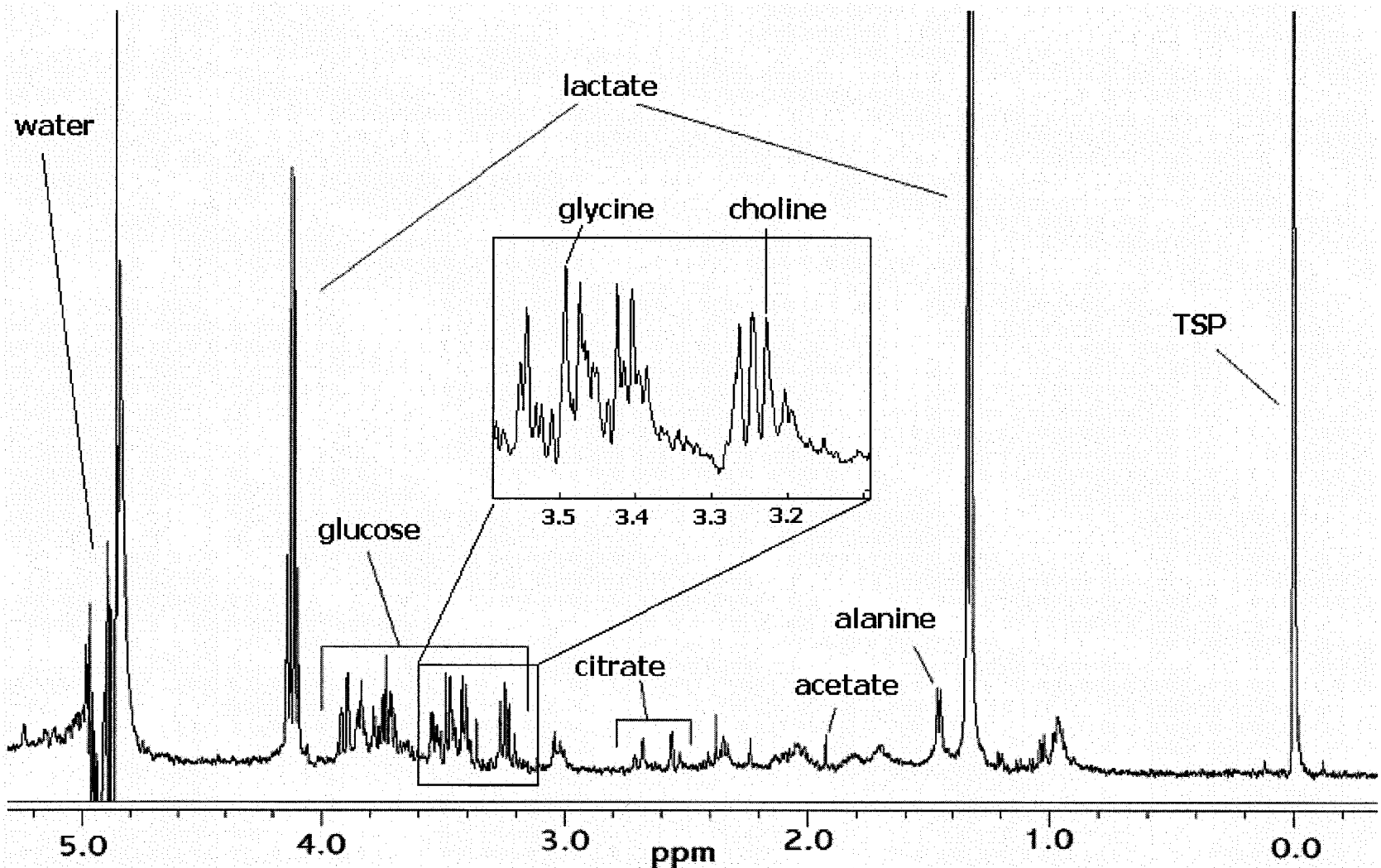

Figure 5.1 – Fourier transform ^{1}H NMR spectrum of 45µL amniotic fluid in 600µL 10% D_2O/90% H_2O from a 500MHz Varian spectrometer

Table 5.1 – Amniotic fluid metabolite chemical shifts in ^{1}H NMR spectrum, with their respective signal multiplicity and functional group

Metabolite	Chemical Shift (Multiplicity)[1]	Function
TSP	0.00	$Si(CH_3)_3$
Lactate	1.33 (d), 4.12 (q)	CH_2, CHOH
Alanine	1.47 (d)	CH_3
Acetate	1.92 (s)	CH_3
Citrate	2. 53 (d), 2. 68 (d)	CH_2, CH_2
Choline	3.24 (s)	$N(CH_3)_3$
Glycine	3.48 (s)	CH_2
α,β-Glucose	3.16-3.96 (various)	CHOH

[1]Chemical shifts are expressed in ppm. The multiplicities are: singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m).

Quantitative Analysis

The quantitative ^{1}H NMR windowed DECRA protocol developed in chapter III was then used on the resulting AF FIDs. The AF metabolites that could be clearly identified were quantified and they include glucose, glycine, choline, citrate, acetate, alanine and lactate. Their windowed frequency regions, which were analyzed by DECRA are listed in table 5.2.

Table 5.2 – Key metabolite frequency regions isolated in the ^{1}H NMR frequency-domain spectra of amniotic fluid for analysis by DECRA

Metabolite	Chemical Shift Isolated Region (ppm)
TSP	-0.150 – 0.150
Lactate	1.260 - 1.400
Alanine	1.412 - 1.511
Acetate	1.915 - 1.934
Citrate	2.484 - 2.749
Choline	3.235 - 3.255
Glucose	3.395 - 3.434
Glycine	3.469 - 3.480

The chemical shift reference, TSP was also quantified to determine metabolite absolute concentrations. Knowing the absolute concentration of TSP (0.14mM), an average value of TSP DECRA concentration estimates for all AF spectra (E_{TSP}) was used as a internal concentration standard for each metabolite concentration ($C_{metabolite}$). Compensating for the number of equivalent hydrogens per signal (TSP n = 9, glucose n = 2, glycine n = 2, choline n = 9, citrate = 4, acetate = 3, alanine = 3, lactate = 2), the absolute concentration of each metabolite was calculated using the internal standard method where,

$$C_{metabolite} = C_{TSP} \times \frac{E_{metabolite}}{E_{TSP}} \times \frac{n_{metabolite}}{n_{TSP}} \qquad 5.1$$

With Equation 5.1, absolute concentrations of AF metabolites can be obtained.

Statistical Analysis

Mean and standard deviations were computed for metabolite concentrations of each of the following groups: patients with GDM and normal patients (Non-GDM).

Variances between the groups were analyzed by the classical t-test and results are reported using the P-value. Metabolites that were significantly different at a 95% confidence interval were defined as predictive biomarkers for GDM.

Notched box and whisker plots were also used to depict the variances between the GDM and non-GDM groupings. The plot contains a central box, whose top and bottom lines limit the lower and upper quartiles. In other words, the lower quartile contains twenty-fifth percentile of the data, and the upper quartile contains the seventy-fifth percentile. The central line of the box plot represents the median, and separates the data set into two equal parts. Lines extending perpendicular from the upper and lower quartile lines, also termed 'whiskers' express the range of the data set. Outliers are data values beyond the limits of the whiskers and are denoted with a '+', and are values that are more than 1.5 times the inter-quartile range (difference between the top and bottom quartile limits).

Notches in a box and whisker plot represent an estimate of the uncertainty about the median, and are useful for box-box comparisons. Specifically, if the notches about two medians do not overlap, the medians are, roughly, significantly different at a confidence level of 95% [78, 79]. Note that the confidence level is based upon a normal distribution assumption. The width of a notch is computed as the inter-quartile range (R) normalized by the square root of the number of data points (N),

$$width = C\left(\frac{1.25R}{1.35\sqrt{N}}\right) \qquad 5.2$$

where C is a constant and is equal to 1.96, a conservative estimate of the variance. Values of C can be chosen to better represent the data, depending on whether the standard

deviations of the two groups are vastly different or similar (refer to references [78, 79] for further details).

5.3 Results

Concentration estimates of lactate, alanine, acetate, citrate, choline, glucose, and glycine were computed using the quantitative ^{1}H NMR windowed DECRA analysis. A summary of these results is given in table 5.3, where the mean concentrations and standard deviations of each metabolite for GDM and Non-GDM groupings are reported. T-test results are also reported in table 5.3, with P-values of less than or equal to 0.05 (α-value) indicate a significant difference of means at a confidence level of 95%.

Table 5.3 – Key amniotic fluid metabolite relative concentration estimates resulting from DECRA analysis for GDM groupings (mean ± STD)

Metabolite	Non-GDM ($x10^{-2}$ mM)	GDM ($x10^{-2}$ mM)	P-value
Lactate	3.2 ± 1.3	3.2 ± 1.1	0.52
Alanine	3.3 ± 1.4	3.2 ± 1.3	0.53
Acetate	0.15 ± 0.04	0.15 ± 0.05	0.34
Citrate	1.5 ± 0.5	1.4 ± 0.5	0.66
Choline	1.3 ± 0.4	1.5 ± 0.5	0.009
Glucose	0.54 ± 0.18	0.66 ± 0.22	0.003
Glycine	0.14 ± 0.05	0.18 ± 0.06	0.004

GDM = patients diagnosed with GDM, Non-GDM = normal patients

T-test results show significant differences between choline, glucose, and glycine concentrations of GDM and Non-GDM groupings, whereas no significant differences are observed for the lactate, alanine, acetate, and citrate mean concentrations.

The notched box and whisker plots representing the metabolite concentration estimates for both GDM and normal (Non-GDM) patients are given in figures 5.2 and 5.3. Two separate figures are used to illustrate the metabolite groups that either change or remain the same between GDM and normal pregnancies, as concluded from the t-test. In figure 5.2, box and whisker plots of AF metabolites with no apparent change in concentrations as a result of GDM have similar median lines with overlapping notches. These traits are especially evident in the box and whisker plots for lactate and alanine concentrations.

Figure 5.3 shows AF metabolite concentrations that significantly differ between diabetic and normal patients. Results from a t-test show that these metabolites are glucose, glycine, and choline. Analyzing the box and whisker plots for these components, little notch overlap between the Non-GDM and GDM groupings is observed. An increase in concentration medians for all three metabolites indicates an increase of these AF constituents in pregnancies with gestational diabetes mellitus.

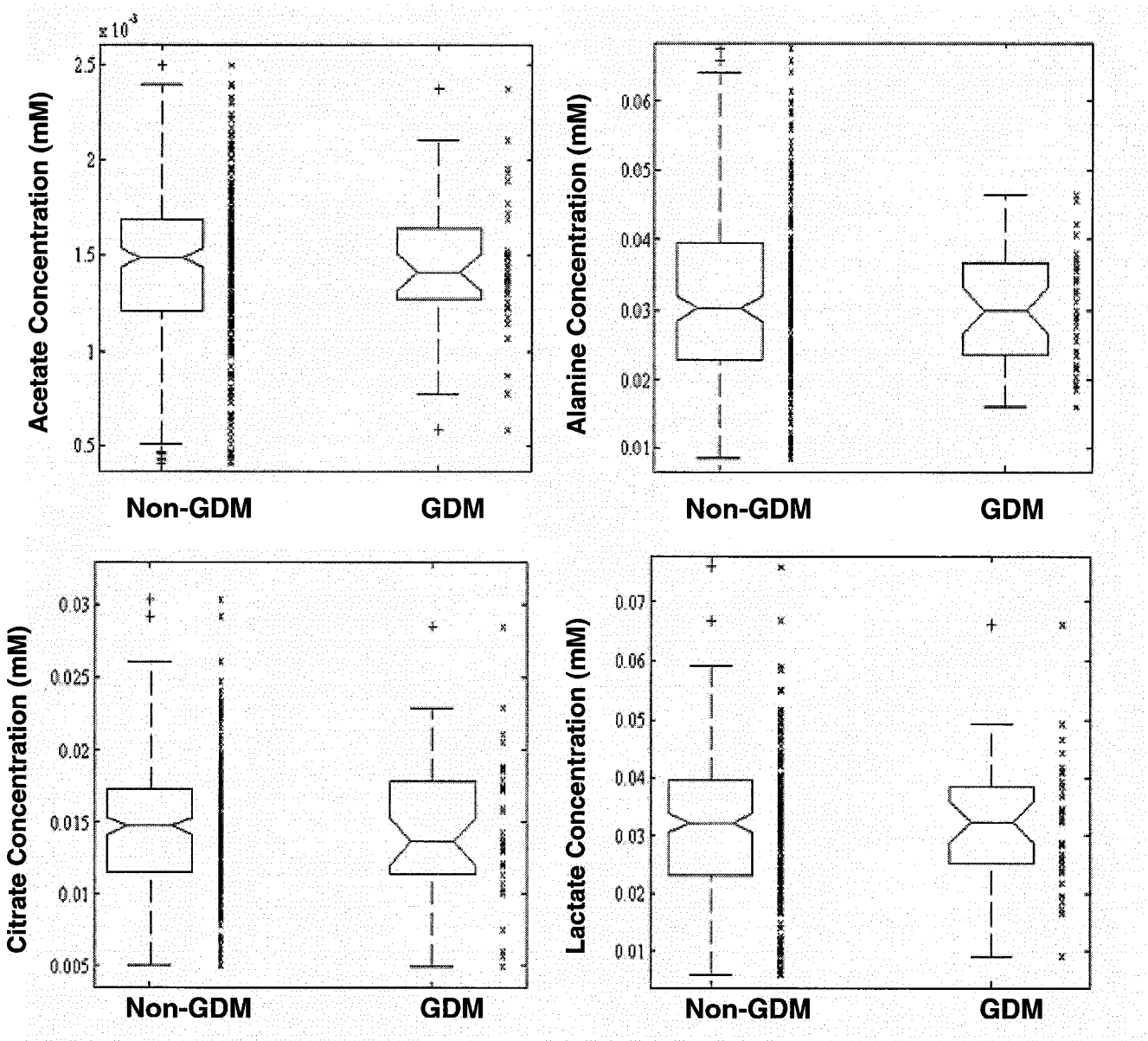

Figure 5.2 – Notched box and whisker plots of GDM and non-GDM patients given alanine, acetate, lactate and citrate concentrations. Lines define the mean, and limit the lower and upper quartiles. Whiskers represent the extent of the data set, and outliers are denoted with a '+'. Data points are given to the right of the box plots.

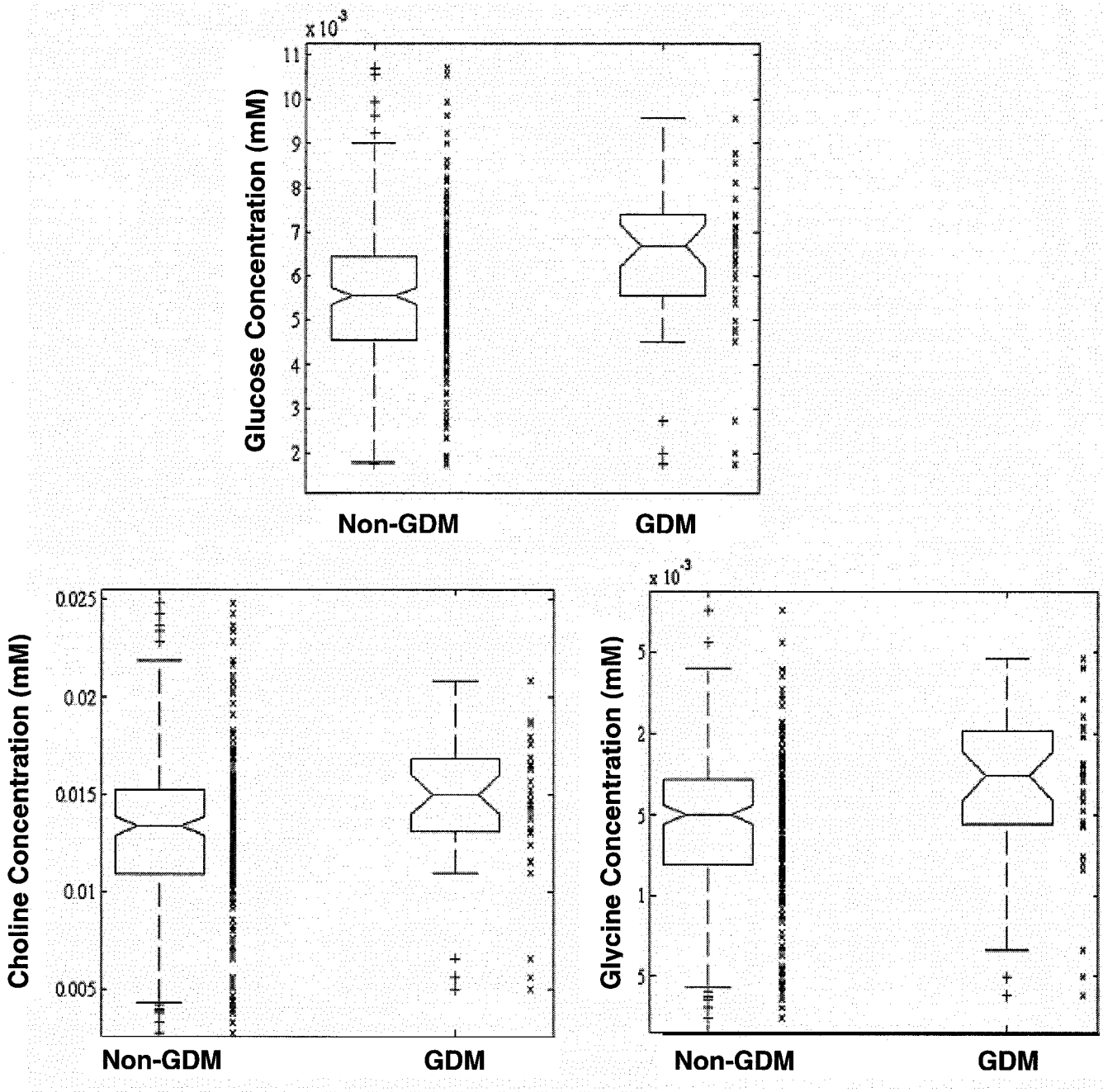

Figure 5.3 – Notched box and whisker plots of GDM and non-GDM patients given glucose, glycine, and choline concentrations. Lines define the mean, and limit the lower and upper quartiles. Whiskers represent the extent of the data set, and outliers are denoted with a '+'. Data points are given to the right of the box plots.

5.4 Discussion

Both alanine and lactate are precursors in the production of glucose in the liver (gluconeogenesis), which is transferred to the blood and muscles. Means for both these AF metabolites concentrations, between GDM and normal pregnancies, are relatively similar. This is consistent with the study by Kalhan et al [60], where the rate of alanine turnover was similar for both diabetic and normal patients during pregnancy, and that glucose production in the liver between the two groupings had no significant difference. As neither acetate, nor citrate is involved in gluconeogenesis, the similar AF concentration means for both GDM and normal patients were expected.

However, glycine has also been reported to be a precursor to gluconeogenesis [80], and a positive relationship was found between glycine and GDM. Alanine, lactate, and glycine have multiple potential sources including the mother, the fetus, and the placenta. Since neither alanine nor lactate concentration increase in diabetic mothers, it can be predicted that glycine originates from an alternate source like the fetus or placenta, as a response to the increased glucose concentration in amniotic fluid.

This positive relationship was also observed between choline and GDM. A phospholipid may be the cause of this observation. Phosphatidylcholine (PC) is synthesized via a CDP-choline pathway and is stimulated with glucose [81]. The increase in choline concentration as a result of GDM can be explained by the increase in glucose stimulant to the PC synthesis. However, it is uncertain if this stimulation occurs in the mother or the fetus. It has also been shown that choline deficiency improved glucose tolerance [82]. The inverse would result in an increase in choline levels as a result of glucose intolerance, as found in diabetic patients.

Previous investigation defines AF glucose as a primary biomarker for determining the onset of GDM. The results of this study in table 5.3 are in agreement with these previous studies, having glucose concentrations in AF increasing in GDM versus normal mothers. This increase in AF glucose is supported by Weiss et al [75], who reported that AF glucose increased by a total of 42% in patients with abnormal glucose tolerance as compared to normal pregnancies. In a similar study, Iyengar et al [73] showed that although AF glucose decreases progressively with advancing gestation, significantly high levels of AF glucose were found in diabetic patients as compared to normal pregnancies. Shiffman et al [74] also reported that elevation in AF glucose concentration, after second trimester amniocentesis, was predictive of subsequent abnormal glucose challenge test, and possibly GDM.

These high glucose levels in GDM patients are thought to be a result of an insulin activity inhibition as a result of hormones produced during pregnancy. Human placental lactogen (hPL), free cortisol, progesterone and prolactin produced during pregnancy lead to the inability of the human body to process glucose [59]. High levels of glucose cross the placental membrane, whereas insulin crossing is inhibited, leading to fetal hyperglycemia.

5.5 Conclusions

This investigation has demonstrated that high-resolution ^{1}H NMR spectra of amniotic fluid (AF) can be useful for metabonomic analysis of gestational disorders. Specifically, metabonomics of AF was useful in the diagnosis of gestational diabetes mellitus (GDM). The study demonstrated that levels of AF glucose, choline, and glycine are relatively good predictive biomarkers and correlate best with this gestational disorder.

Simultaneous metabolic profiling and quantitative analysis was possible with the application of the windowed DECRA method. One advantage to this method is that a pattern of several metabolites is more informative than measurement of any single analyte by a traditional biochemical assay. Qualitative and quantitative information is contained within a single NMR spectrum and numerous metabolites can be quantified simultaneously.

The majority of previous metabonomic studies of amniotic fluid in relation to GDM were done during second trimester or late periods of gestation. In this investigation, the same conclusions were drawn using AF obtained earlier on in pregnancy. However, when computing the percent difference between GDM and normal AF glucose concentrations, a value of 1.2% is obtained, as compared to 42% given in the study by Weiss et al. The difference between these two results can be explained by the average gestational week in which AF specimens were collected. Weiss et al's study compared glucose intolerant AF samples collected at an average of 33.9 weeks gestation, as compared to an average of 15.3 weeks in the present study. Notably, no AF was collected before 14 weeks gestation in Weiss's study, and no comparison measurements were made before 28 weeks. Although a significant difference is observed earlier on in pregnancy, the sensitivity of the diagnosis decreases. Nevertheless, the present study's results promote quantitative NMR as a potential tool for earlier diagnosis of GDM in pregnancy.

VI. General Conclusions

This thesis demonstrates a protocol for high-resolution quantitative proton nuclear magnetic resonance (^{1}H NMR) of biological fluids. Using the windowed direct exponential curve resolution algorithm (DECRA), accurate quantification of biological solutions was achieved. Applying the method to time-domain NMR signals (FIDs) of amniotic fluid, key metabolites in the complex biological fluid were quantified with high specificity.

FIDs were acquired by optimizing NMR parameters such as digital resolution, pulse widths, relaxation delay times, signal to noise ratio, and appropriate post-acquisition processes. Optimization resulted in an NMR signal that could be quantifiable, and obey the relationship given by Equation 2.1. The sensitivity of DECRA to instrumental effects was tested using simulated FIDs, and the major sources of error contributing to the DECRA accuracy were determined. These limitations included baseline distortions and additions, and variable peak line-widths. The DECRA method was calibrated by adjusting these factors for a real data set and calibration biological solutions. An improved accuracy of the quantitative analysis method resulted by zeroing signal baselines, and Lorentzian apodizations to smooth signal line-widths. Calibration resulted in coefficient of variances of 3.3% and 3.7% for alanine and glucose concentrations.

The windowed DECRA analysis method, however accurate, still possesses limitations. DECRA fails when more than approximately four components are contained within a windowed region in an NMR spectrum [42]. For highly complex signals, like biological fluids, many subsections must be isolated to compute concentration for all

metabolites signals visible. This is process could become tedious and time-consuming, and could limit the method's application to simple solutions.

Another limitation that may be overlooked is the Nuclear Overhauser Effect (NOE) as a result of solvent suppression. Irradiation of the water signal can affect the neighboring spin state populations in such a way as to increase the intensity of their respective signals. The NOE interaction between two proton magnetic dipoles takes place predominantly through space interactions as opposed to bond interactions (polarization transfer). Given that the water molecules are in constant motion throughout the sample, it would be difficult to determine the exact affects of the NOE in the sample spectrum. However multiple two-dimensional homonuclear NMR experiments, such as NOESY, may be useful in overcoming this limitation.

Slight errors also arise in the quantitative relationship (Equation 2.1) when setting the relaxation delay time (*dt*). Recall that setting *dt* to $5\times T_1$ (max) only allows for 99.3% of the total signal intensity. Meaning the slowest relaxing nuclei will only have a maximum signal intensity of 99.3%, giving slightly skewed relative amounts, unreflective of metabolites within the sample. To overcome this limitation, delay times can be set to longer times than $5\times T_1$ (max). However, when making this change, consideration to the number of samples, and to the total acquisition time should be kept in mind. As this current study required 258 consecutive FIDs to be collected, longer acquisition times would have been inconvenient.

The windowed DECRA quantitative protocol outlined can be applied to metabonomic studies of numerous biological fluids in response to pathological and/or physiological conditions. In this study, key amniotic fluid metabolites quantified by the

protocol outlined include glucose, glycine, choline, acetate, alanine, citrate, and lactate. These AF metabolite concentrations were grouped according to the corresponding patients diagnosis of gestational diabetes mellitus. Means and standard deviations of these resulting groups (GDM, N=35; non-GDM, N=223) were computed and compared using the statistical t-test, and notched box and whisker plots. According to the t-test the significant biomarkers for gestational diabetes mellitus, from the metabolites investigated, were found to be glucose (P = 0.004), glycine (P = 0.003), and choline (P = 0.009). Box and whisker plots demonstrated little to no overlap between notches, representing rough standard deviations about the medians, for each grouping.

This study proved a powerful analytical tool for the diagnosis of gestational diabetes mellitus at earlier stages of pregnancy.

Future Work

As the results in the present study are preliminary, improvements could be made to multiple portions, in order to investigate other instrumental effects and limitations to the DECRA sensitivity. Particularly, calibration of DECRA using biological solutions in chapter III was prepared using a serial dilution method. However, if the constituents independently varied, interference from overlapping signals could be examined. Also, larger windowed regions, encompassing additional components can be explored to determine the limitations to the number of components resolvable by DECRA.

Other biological fluids can be analyzed with the quantitative ^{1}H NMR protocol developed. For example, quantification of constituents in *in vitro* fertilization fluid, seminal plasma, and follicular fluid can be studied using this analytical technique to be

correlated with pregnancy outcomes. Urine and blood plasma are also interesting candidates to be analyzed by the quantitative protocol.

An interesting application of quantitative ^{1}H NMR analysis is its potential use as an *in vivo* diagnostic tool. Given the results of the proton NMR metabonomics of amniotic fluid in relation to gestational diabetes mellitus, a future application of *in vivo* maternal health diagnosis might be possible. MRS, the application of proton NMR spectroscopy to *in vivo* measurements, is continually improving on its resolution capabilities, making it possible to apply the DECRA method at high resolutions. However, several countries have issued restrictions on the maximum specific absorption rate (SAR, W/kg) of the human body, limiting magnetic field strengths of MRS and MRI systems. There is a positive correlation with power of the RF pulse used to excite proton spins within the human body, and body temperature. Increasing magnetic field strength, increases the risk of hyperthermia [83]. SAR limits issued by the FDA include 4 W/kg averaged over the whole body for any 15-minute period, 3 W/kg averaged over the head for any 10-minute period, and 8 W/kg in any gram of tissue in the extremities for any period of 5 minutes[83]. Yet, recent studies have shown increased resolution to MR images and spectra, of both the human brain and mouse brain, at lower magnetic field strengths [84, 85]. These advances incorporate fast spin echo (FSE) pulse sequence, commonly employed for the acquisition of T_2 weighted images at magnetic field strengths between 1.5 – 4.7T (NOTE: conversion of T (tesla) to the Larmor frequency (MHz) incorporates the nuclear gyromagnetic ratio, γ). With these progressive improvements, future applications of the quantitative ^{1}H NMR protocol developed

combined with MRS could be a powerful diagnostic tool in clinical community for, not only maternal health during pregnancy, but for many other metabolic disorders.

In addition to potential *in vivo* applications, the protocol developed for quantitative high-resolution ^{1}H NMR spectroscopy has numerous applications to quantifying numerous biological fluids, advancing the field of metabonomics. NMR spectroscopy, using the windowed DECRA method, is no longer limited to only qualitative studies of biological systems, but can be used to produce quantitative metabolic profiles of any biological system.

REFERENCES

1. Lindon, J., *Metabonomics – Techniques and Applications*. 2004, Metabometrix Ltd. p. 1-6.

2. Guido F. Pauli, B.U.J., David C. Lankin, *Quantitative 1H NMR: Development and Potential of a Method for Natural Products Analysis.* Journal of Natural Products, 2004. **68**(1): p. 133-149.

3. Alison Nordon, P.J.G., Colin A McGill, and David Littlejohn, *Quantitative Analysis of Low-Field NMR Signals in the Time Domain.* Analytical Chemistry, 2001. **73**(No. 17): p. 4286-4294.

4. Macomber, R.S., *A Complete Introduction To Modern NMR Spectroscopy*. 1998: John Wiley & Sons, Inc.

5. Wuthrich, K., *NMR of Proteins and Nucleic Acids*. 1986, New York: Wiley.

6. Wüthrich, K., *Protein recognition by NMR.* Nature Structural Biology, 2000. **7**(3): p. 188-189.

7. R. Kaptein, R.B., R. M. Scheek, W. F. van Gunstered, *Protein Structures From NMR.* Perspectives In Biochemistry, 1988. **27**(15): p. 5389-5395.

8. F. Maltz, H.J., *Validation of Quantitative NMR.* Journal of Pharmaceutical and Biomedical Analysis, 2005. **38**: p. 813-823.

9. Csaba Szantay, J., *High-field NMR Spectroscopy As An Analytical Tool For Quantitative Determinations: Pitfalls, Limitations, and Possible Solutions.* Trends In Analytical Chemistry, 1992. **11**(9): p. 332-344.

10. Ray, J., *qNMR Experiments: Key Parameters for Quantitation*. 2004-2006.

11. David J. Cookson, B.E.S., *Optimal Experimental Parameters for Quantitative Pulse Fourier Transform Proton Nuclear Magnetic Resonance Spectroscopy.* Analytical Chemistry, 1982. **54**: p. 2591-2593.

12. U. Holzgrabe, R.D., C. Schollmayer, B. Waibel, *Quantitative NMR Spectroscopy - Applications in Drug Analysis.* Journal of Pharmaceutical and Biomedical Analysis, 2005. **38**: p. 806-812.

13. P.E. McGowan, J.R., R. Wilson, J.J. Walker, S. Wisdom, J.H. McKillop, *Quantitative 1H-NMR Analysis of Amniotic Fluid.* Journal of Pharmaceutical and Biomedical Analysis, 1993. **11**(8): p. 629-632.

14. Bock, J.L., *Metabolic Profiling of Amniotic Fluid by Proton Nuclear Magnetic Resonance Spectroscopy: Correlation with Fetal Maturation and Other Clinical Variables.* Clinical Chemistry, 1994. **40**(1): p. 56-61.

15. P.E. McGowan, W.C.L., J. Reglinski, C.M. Spickett, R. Wilson, J.J. Walker, S. Wisdom, M.A. Maclean, *1H NMR As A Non-Invasive Probe of Amniotic Fluid in Insulin Dependent Diabetes Mellitus.* Journal of Perinatal Medicine, 1999. **27**: p. 404-408.

16. Ian W. Burton, M.A.Q., John A. Walter, *Quantitative 1H NMR with External Standards: Use in Preparation of Calibration Solutions for Algal Toxins and Other Natural Products.* Analytical Chemistry, 2005. **77**: p. 3123-3131.

17. D. J. Crockford, H.C.K., L. M. Smith, E. Holmes, J. K. Nicholson, *Curve-Fitting Method for Direct Quantitation of Compounds in Complex Biological Mixtures Using 1H NMR: Application in Metabonomic Toxicology Studies.* Analytical Chemistry, 2005.

18. Provencher, S.W., *Estimation of Metabolite Concentrations From Localized In Vivo Proton NMR Spectra.* Magnetic Resonance Medicine, 1993. **30**(6): p. 672-679.

19. Martin, Y.L., *A Global Approach to Accurate and Automatic Quantitative Analysis of NMR Spectra by Complex Least-Squares Curve Fitting.* Journal of Magnetic Resonance, 1994. **111**(1): p. 1-10.

20. Leentie Vanhamme, T.S., Paul Van Hecke, Sabine Van Huffel, *MR Spectroscopy Quantitation: A Review of Time-Domain Methods.* NMR in Biomedicine, 2001. **14**: p. 233-246.

21. J.W.C. van der Veen, R.d.B., P.R. Luyten, D. van Ormondt, *Accurate Quantification of In Vivo 31P NMR Signals Using The Variable Projection Method and Prior Knowledge.* Magnetic Resonance Medicine, 1988. **6**: p. 92-98.

22. Leentie Vanhamme, A.v.d.B., S. Van Huffel, *Improved Method For Accurate and Efficient Quantification of MRS Data With Use of Prior Knowledge.* Journal of Magnetic Resonance, 1997. **129**: p. 35-43.

23. Brian Antalek, W.W., *Generalized Rank Annihilation Method Applied to a Single Multicomponent Pulsed Gradient Spin Echo NMR Data Set.* Journal of American Chemical Society, 1996. **118**: p. 10331-10332.

24. Bruce Wilson, E.S., Bruce R. Kowalski, *An Improved Algorithm For The Generalized Rank Annihilation Method.* Journal of Chemometrics, 1989. **3**: p. 493-498.

25. E. Sanchez, B.R.K., *Generalized Rank Annihilation Factor Analysis.* Analytical Chemistry, 1986(58): p. 469-499.

26. Willem Windig, B.A., *Direct Exponential Curve Resolution Algorithm (DECRA): A Novel Application of The Generalized Rank Annihilation Method For A Single Spectral Mixture Data Set With Exponentially Decaying Contribution Profiles.* Chemometrics and intelligent laboratory systems, 1997. **37**: p. 241-254.

27. Alison Nordon, C.M., Robert H. Carr, Paul J. Gemperline, David Littlejohn, *Determination of The Ethylene Oxide Content of Polyether Polyols By Low-field 1H Nuclear Magnetic Resonance Spectrometry.* Analytica Chimica Acta, 2002. **472**: p. 133-140.

28. Kan-Zhi Liu, M., PhD, and Henry H. Mantsch, PhD, *Simultaneous Quantitation From Infrared Spectra of Glucose Concentrations, Lactate Concentrations, and Lecithin/Sphingomyelin Ratios In Amniotic Fluid.* American Journal of Obstetrics and Gynecology, 1999. **180**(3): p. 696-702.

29. Teresa M. Rudkin, D.L.A., *Proton Magnetic Resonance Spectroscopy For The Diagnosis of Cerebral Disorders.* Arch. Neurol., 1999. **56**: p. 919-926.

30. Damien Galanaud, F.N., Olivier Chinot, Sylviane Confort-Gouny,Dominique Figarella-Branger, Pierre Roche, Stephane Fuentes, Yann Le Fur, Jean-Philippe Ranjeva, Patrick J. Cozzone, *Noninvasive Diagnostic Assessment of Brain Tumors Using Combined In Vivo MR Imaging and Spectroscopy.* Magnetic Resonance in Medicine, 2006. **55**: p. 1236-1245.

31. M. Filippi, V.D., H.F. McFarland, D.H. Miller, R.I. Grossman, *Role of Magnetic Resonance Imaging in the Diagnosis and Monitoring of Multiple Sclerosis: Consensus Report of the White Matter Study Group.* Journal of Magnetic Resonance Imaging, 2002. **15**: p. 499-504.

32. *Magnetic Resonance Medicine: Official Journal of the Society of Magnetic Resonance in Medicine*, United States: Wiley-Liss, Inc.

33. *Magnetic Resonance in Medical Sciences: MRMS: An Official Journal of Japan Society of Magnetic Resonance In Medicine*, Japan: Japan Soceity of Magnetic Resonance in Medicine.

34. T.R. Nelson, R.J.G., D. A. Powell, M.C. Schrader, D.K. Manchester, D. H. Pretorius, *High Resolution Proton NMR Spectroscopy of Human Amniotic Fluid.* Prenatal Diagnosis, 1987. **7**: p. 363-372.

35. Rolf Gruetter, S.A.W., Vasantham Rajanayagan, Melissa Terpstra, Hellmut Merkle, Charles L. Truwit, Michael Garwood, Scott L. Nyberg, Kamil Ugurbil, *Resolution Improvements in in Vivo 1H NMR Spectra with Increased Magnetic Field Strength.* Journal of Magnetic Resonance, 1998. **135**: p. 260-264.

36. Tsang-Lin Hwang, A.J.S., *Water Suppression That Works. Excitation Sculpting Using Arbitrary Waveforms and Pulsed Field Gradients.* Journal of Magnetic Resonance, 1995. **Series A 112**: p. 275-279.

37. Ieda Scarminio, M.K., *Analysis of Correlated Spectral Data.* Analytical Chemistry, 1993. **65**: p. 409-416.

38. David H. Burns, J.B.C., Gary D. Christian, *Rank Annihilation with Incomplete Information.* Analytical Chemistry, 1986. **58**: p. 2805-2811.

39. E. Sanchez, B.R.K., *Generalized Rank Annihilation Factor Analysis.* Analytical Chemistry, 1986. **58**: p. 496-499.

40. K.S. Booksh, B.R.K., *Comments On: The Data Analysis (Datan) Algorithm and Rank Annihilation Factor Analysis For The Analysis Of Correlated Spectral Data.* Journal of Chemometrics, 1994. **8**(4): p. 287-292.

41. Silbs, P., *Fourier Transform Pulsed-Gradient Spin-Echo Studies of Molecular Diffusion.* Progress In NMR Spectroscopy, 1987. **19**: p. 1-45.

42. Antalek, B., *Using Pulse Gradient Spin Echo NMR for Chemical Mixture Analysis: How to Obtain Optimum Results.* Concepts in Magnetic Resonance, 2002. **14**(4): p. 225-258.

43. Plentl, A.A., *Formation And Circulation Of Amniotic Fluid.* Clinical Obstetrics and Gynecology, 1966. **9**(2): p. 427-439.

44. Ostergard, D.R., *The physiology and clinical importance of amniotic fluid. A review.* Obstetrical & Gynecological Survey, 1970. **25**(4): p. 297-319.

45. *Amniotic Fluid: Research and Clinical Application.* 2nd ed, ed. T.K.A.B.E. D.V.I Fairweather. 1978, Netherlands: Elsevier/North-Holland Biomedical Press.

46. *Amniotic Fluid And Its Clinical Significance*, ed. M. Sandler. 1981, New York: Marcel Dekker, Inc.

47. Barth H. Ragatz, G.M., *Estimation of Amniotic Fluid Phospholipids Assessing Fetal Lung Maturity.* Biochemical Education, 1985. **13**(3): p. 134-136.

48. P. Rosati, P.P., P. Riccardi, R. Flore, P. Tondi, U. Bellati, *The Use Of Amniotic Fluid Viscosity Measurements To Establish Fetal Lung Maturity.* International Journal of Gynecology and Obstetrics, 1991. **35**: p. 351-355.

49. Bradford W. Fenton, C.-S.L., Frank Seydel, Christian Macedonia, *Lecithin Can Be Detected By volume-Selected Proton MR Spectroscopy Using a 1.5T Whole Body Scanner: A Potentially Non-Invasive Method For The Prenatal Assessment of Fetal Lung Maturity.* Prenatal Diagnosis, 1998. **18**: p. 1263-1266.

50. Anne-Lise Bjbrke Monsena, J.S., Per Magne Uelanda, *Mid-trimester Amniotic Fluid Methionine Concentrations: A Predictor Of Birth Weight and Length.* Metabolism Clinical and Experimental, 2006. **55**: p. 1186-1191.

51. C.D. Naylor, M.S., E. Chen, K. Sykora, *Cesarean Delivery In Relation To Birth Weight and Gestational Glucose Tolerance: Pathophysiology or Practice Style? Toronto Trihospital Gestaional Diabetes Investigators.* The Journal of The American Medical Association, 1996. **275**(15): p. 1165-1170.

52. J. Verhaeghe, R.V.B., E. Van Herck, J. Laureys, R. Bouillon, F.A. Van Assche, *C-peptide, insulin-like growth factors I and II, and insulin-like growth factor binding protein-1 in umbilical cord serum: correlations with birth weight.* American Journal of Obstetrics and Gynecology, 1993. **169**(1): p. 89-97.

53. Marshall W. Carpenter, J.A.C., Joseph W. Hogan, Curtis Shellum, Margaret Somers, Jami A. Star, *Amniotic Fluid Insulin at 14–20 Weeks' Gestation: Association with later maternal glucose intolerance and birth macrosomia.* Diabetes Care, 2001. **24**(7): p. 1259-1263.

54. Pamela J. Surkan, C.-C.H., Anna L.V. Johansson, Paul W. Dickman, Sven Cnattingius, *Reasons for Increasing Trends in Large for Gestational Age Births.* The American College of Obstetricians and Gynecologists, 2004. **104**(4): p. 720-726.

55. Pederson, J., *Weight and Length At Birth of Infants of Diabetic Mothers.* Acta Endocrinol, 1954. **16**: p. 330-342.

56. Prakeshkumar Shah, A.O., *Literature Review of Low Birth Weight, Including Small for Gestational Age and Preterm Birth.* 2002, Toronto Public Health: Toronto. p. 1-134.

57. Theodor Stefos, A.S., Apostolos Kaponis, Nicolaos Dalkalitsis, Dimitrios Lolis, *Amniotic Fluid Glucose At The Time of Genetic Amniocentesis: Correlation With Duration Of Pregnancy and Birthweight.* European Journal of Obstetrics & Gynecology and Reproductive Biology, 2003. **106**: p. 144-147.

58. Siri L. Kjos, T.A.B., *Gestational Diabetes Mellitus.* New England Journal of Medicine, 1999. **341**(23): p. 1749-1756.

59. Jean-Louis Chiasson, G.G.E.A., Francine Ducros, Josee Bourque, Pierre Maheux, *Glucose Turnover and Gluconeogenesis During Pregnancy In Women With and Without Insulin-Dependent Diabetes Mellitus.* Clinical and Investigative Medicine, 1997. **20**(3): p. 140-151.

60. SC. Kalhan, R.H.H., K.Q. Rossi, S.M. Savin, *Glucose-Alanine Relationship in Diabetes in Human Pregnancy.* Metabolism, 1991. **40**(6): p. 629-633.

61. Jami Star, J.A.C., Glenn E. Palomaki, Marshall W. Carpenter, Devereux N. Sallier Jr., C. James Sung, Marea B Tumber, Donald R. Coustan, *The Relationship Between Second-Trimester Amniotic Fluid Insulin and Glucose Levels and Subsequent Gestational Diabetes.* Prenatal Diagnosis, 1997. **17**(2): p. 149-154.

62. Emanuela Meddaa, S.D., Angela Spinellia, Gian Carlo Di Renzob, *Genetic Amniocentesis: A Risk Factor For Preterm Delivery?* European Journal of Obstetrics & Gynecology and Reproductive Biology, 2003. **110**: p. 153–158.

63. Seeds, J.W., *Diagnostic mid trimester amniocentesis: How safe?* American Journal of Obstetrics and Gynecology, 2004. **191**: p. 608-616.

64. M. Cederholm, B.H., O. Axelsson, *Amniocentesis At 14-15 Weeks Gestation Increased Risk Of Infant Postural Deformities And Respiratory Distress.* Evidence-based Obstetrics and Gynecology, 2006. **8**: p. 7-8.

65. Eric Jauniaux, G.S.P., Charles H. Rodec, *What Invasive Procedure To Use In Early Pregnancy?* **Bailliere's Clinical Obstetrics and Gynaecology**, 2000. **14**(4): p. 651-662.

66. Hardy, D.S., *A Multiethnic Study of Predictors of Macrosomia.* The Diabetic Educator, 1999. **25**(6): p. 925-933.

67. Jovanovic, L., *The Role of Continuous Glucose Monitoring in Gestational Diabetes Mellitus.* Diabetes Technology, 2000. **2**: p. S-67 - S-71.

68. *Mosby's Manual of Diagnostic and Laboratory Tests.* 2 ed, ed. T.J.P. K.D. Pagana. 2002, St. Louis: Mosby.

69. *Manual of Laboratory And Diagnostic Tests.* 7 ed, ed. M.B.D.I. F. T. Fischbach. 2004, Philadelphia: Lippincott Williams and Wilkins.

70. *Laboratory Tests and Diagnostic Procedures.* 4 ed, ed. B.J.B. C.C. Chernecky. 2004, Philadelphia: Saunders.

71. . *Handbook of Diagnostic Tests.* 3 ed. 2003, Philadelphia: Lippincott Williams and Wilkins.

72. Endocrinologists, A.A.o.C., *Position Statement On Metabolic And Cardiovascular Consequences Of Polycystic Ovary Syndrome.* Endocrine Practice, 2005. **11**(2): p. 126-134.

73. Leela R. Iyengar, V.A.R., S. Kumar, *Significance of Amniotic Fluid Glucose In Pregnancy.* International Federation of Gynaecology and Obstetrics, 1982. **20**: p. 57-63.

74. R.L. Shiffman, D.F., G.S. Johnson, E. Schutta, M. Wolf, F. Mejia, *Amniotic Fluid Glucose and Gestational Diabetes.* American Journal of Obstetrics and Gynecology, 1991. **164**: p. 286.

75. P. A. Weiss, H.H., R. Winter, P Purstner, W. Lichtenegger, *Amniotic Fluid Glucose values In Normal and Abnormal Pregnancies.* Obstetrics & Gynecology, 1985. **65**: p. 333-339.

76. Pascal M.W. Groenen, U.F.E., Ron A. Wevers, Jan C.M. Hendriks, Tom K.A.B Eskes, Hans M.W.M. Merkus, Regine P.M. Steegers-Theunissen, *High-resolution 1H NMR Spectroscopy of Amniotic Fluids From Spina Bifida Fetuses and Controls.* European Journal of Obstetrics and Gynecology and Reproductive Biology, 2004. **112**(16-23).

77. Laurence Le Moyec, F.M., Michel Eugene, Manfred Spraul, *Proton Magnetic Resonance Spectroscopy of Human Amniotic Fluids Sampled at 17-18 Weeks of*

Pregnancy in Cases of Decreased Digestive Enzyme Activities and Detected Cystic Fibrosis. Clinical Biochemistry, 1994. **27**(6): p. 475-483.

78. Nelson, L.S., *Evaluating Overlapping Confidence Intervals.* Technometrics, 1989. **21**: p. 140-141.

79. Robert McGill, J.W.T., Wayne A. Larsen, *Variations of Box Plots.* The American Statistician, 1978. **32**(1): p. 12-16.

80. Geza Hetenyi, J., Peter J. Anderson, Mani Raman, Catherine Ferrarotto, *Gluconeogenesis from glycine and serine in fasted normal and diabetic rats.* Biochem. J., 1988. **253**(27-32).

81. J.M. Hoffman, S.G.L., *Choline turnover in phosphatidylcholine of pancreatic islets. Implications for CDP-choline pathway.* Diabetes, 1988. **37**(11): p. 1489-1498.

82. Peter J. Raubenheimer, M.J.N., Brian R. Walker, *A Choline-Deficient Diet Exacerbates Fatty Liver but Attenuates Insulin Resistance and Glucose Intolerance in Mice Fed a High-Fat Diet.* Diabetes, 2006. **55**: p. 2015-2020.

83. Portal, M.R.-T.I., *Specific Absorption Rate, http://www.mr-tip.com/.* Copyright © 2003 - 2006.

84. O. Natt, T.W., S. Boretius, J. Radulovic, J. Frahm, T. Michaelis, *High-resolution 3D MRI of mouse brain reveals small cerebral structures in vivo.* Journal of Neuroscience Methods, 2002. **120**(2): p. 203-209.

85. E. De Vita, D.L.T., S. Roberts, H.G. Parkes, R. Turner, P. Kinchesh, K. Shmueli, T.A. Yousry, R.J. Ordidge, *High resolution MRI of the brain at 4.7 Tesla using fast spin echo imaging.* British Journal of Radiology, 2003. **76**(909): p. 631-637.

APPENDIX A

```
% DECRA (Direct Exponential Curve Resolution Algorithm) analysis on input FID
% Jan 24, 2006

% Make a Hankel Matrix of the spectrum with ncom (number of components rows
data1=zeros(endspec,endspec);
points=endspec;
for i=1:endspec
   data1(i,1:points)=data(i:endspec);
   points=points-1;
end;

%define two ranges (split the data set in two)
range1=[startspec:endspec-1];
range2=[startspec+1:endspec];

%Create a common base for the two data sets using SVD(to obtain square matrices,
%uc and vc.)
[vc,sc,uc]=svd(data1(range1,:)',0); %economy sized svd
sc=sc(1:ncom,1:ncom); % --> Solved by making data into more than a row vector
uc=uc(:,1:ncom);     % time axis (or sample axis) (v and u are the reverse)
vc=vc(:,1:ncom);     % sample axis (or time axis)

%Project the two data sets onto the common base. auv and buv are both square
%matricies of the two data sets. Need square matrices to solve the eigenproblem
auv=sc;
buv=uc'*data1(range2,:)*vc;

%solve the generalized eigenvalue problem: AZa=BZ, where Z=(P')^-1, or the
%pseudoinverse of the pure spectra
[v,s]=eig(buv*inv(sc)); %v=eigenvectors, s=eigenvalues in diagonal matrix
ev=diag(s);
[ev,sortindex]=sort(diag(s));
v=v(:,sortindex); % sort composition in second set of data

%calculate spectra and concentrations
pspec=pinv(vc*inv(sc)*v);
pint=uc*v;
```

APPENDIX B

Refer to the following documents for Ethics approval forms received from McGill University for the study entitled "The Role of Amniotic Fluid Constituents in Predicting Human Fetal Growth".