Université d'Ottawa • University of Ottawa

Applications of the *In Vacuo* Chemical Modification Technique to the Study of Protein Structure and Function

Helen T. Vakos

ABSTRACT

Applications of the *In Vacuo* Chemical Modification Technique to the Study of Protein Structure and Function

Helen T. Vakos

Sensitive NMR methods have been developed for analyzing N- and C-termini in protein structure for investigating protein purity, identity, and homogeneity. The *in vacuo* chemical modification procedure[1,2] allows for the incorporation of isotopically labeled reagents (^{13}C, ^{14}C, ^{2}H, ^{3}H) into protein with a high degree of reaction specificity in a cost-effective manner, since there is no competing reaction with water and only microlitre quantities are required to derivatize milligram quantities of lyophilized protein. The pH memory effect[3] is used to control the ionization state of functional groups in lyophilized proteins, and hence, their chemical reactivity towards modifying reagents.

Reaction *in vacuo* at 75°C with ^{13}C-iodomethane results in the preferential ^{13}C-trimethylation of α-amino groups in proteins lyophilized from aqueous solutions at pH (LpH) 6 to 7. *In vacuo* trimethylation of α-amino groups and analysis by ^{13}C-NMR spectroscopy provides a new approach for determining the *number and types* of N-terminal amino acid residues and for verifying the presence of a *free or blocked* amino-terminus in a protein preparation. Minor resonances of free α-amino groups arising from proteolysis of the major protein or from protein impurities are detectable. SDS-PAGE cannot distinguish subtle differences in contaminating proteins of similar size or heterogeneity. *In vacuo* reaction with ^{13}C-iodomethane also results in the formation of ^{13}C-methyl esters of protein carboxyl groups, the resonances of which must be removed by hydroxylamine or base and heat to clarify the spectra. ^{13}C- and ^{14}N-NMR spectroscopy are utilized to characterize the trimethylammonium derivatives of α- and ε-amino groups formed by reaction with iodomethane (^{13}C, ^{12}C).

Reaction of lyophilized proteins *in vacuo* at 75°C with ^{13}C-iodomethane at LpH values between 3 and 4 effects the preferential esterification of carboxyl groups, and is in accord with the pH memory effect[3]. The extent of reaction increases with an increase in amount of deprotonated carboxyl groups at higher LpH values, but trimethylation of amino groups is the predominant reaction at LpH values greater than 4 or 5. In contrast, the novel *in vacuo* esterification reaction (75°C) of lyophilized proteins (LpH 3) with gaseous ^{13}C-methanol/HCl or ^{13}C-ethanol/HCl results in the exclusive, rapid, and extensive formation of α-, γ- and δ-carboxyl ^{13}C-methyl or ^{13}C-ethyl ester derivatives with no protein degradation, which could be distinguished by the distinct chemical shifts of their resonances. Within the class of γ- or δ-carboxyl esters, the chemical shifts have little variation, however, the chemical shift of a C-terminal α-carboxyl group shows a strong dependence on the nature of the C-terminal amino acid and sequence. Esterification and analysis by ^{13}C-NMR spectroscopy provides a new approach for determining the *number* of *different* C-termini present in a protein preparation. The extent of derivatization of protein carboxyl groups can be controlled and exploited in structure-function studies, as the protein can retain its structural integrity in the lyophilized state throughout the *in vacuo* procedure.

Reaction *in vacuo* with ^{14}C- or ^{3}H-enriched iodomethane results in a higher incorporation of specific radioactivity into lyophilized proteins than can be achieved in aqueous environments. The chemical tag can serve as a probe of structure and function in tracer studies for proteins that are not enzymes. Trace-radiolabeled proteins (e.g., ^{14}C-methylated *Bacillus thuringiensis* (*Bt*) insecticidal toxin, ^{14}C-methylated insulin) show a high specific activity, and retain bioactivity (e.g., ^{14}C-methylated *Bt* toxin).

In a separate study, pure bovine and porcine cholesterol esterase (CE) and human milk bile salt-activated lipase are obtained using a cholate-derivatized affinity column and a taurocholate concentration gradient in the eluent, which is critical to the success of the method. Taurocholate induces dimerization of bovine CE in SDS as observed by SDS-PAGE. Studies of the bile salt-modulated stereoselectivity of CE-catalyzed hydrolysis of α-tocopheryl acetates show that, in competitive experiments, the diastereoselectivity of CE depends on the bile salt used, which is relevant if the bile salt-modulated effect on CE-catalyzed reactions is to be exploited in organic syntheses. A mechanistic study with synthetic α-tocopheryl ester substrates of varying chain length shows that the effect of the latter on CE activity reflects the lifetime of the acylenzyme intermediates and product inhibition. Pentanoate ester is hydrolyzed at a faster rate cf. acetate and decanoate esters. The results are consistent with a rate-limiting deacylation stage of the CE mechanism for the hydrolysis reaction.

1. Taralp, A. (1997) In *Aqueous and Nonaqueous Chemical Approaches to Elucidating Structure and Function in Proteins.* Ph.D. Thesis, Department of Chemistry, University of Ottawa, Ottawa, ON.

2. Taralp, A. and Kaplan, H. (1997) *J. Protein Chem.* **16:** 183-193.

3. Zaks, A. and Klibanov, A. M. (1988) *J. Biol. Chem.* **263:** 3194-3201.

To my family, my father, Theodore, my mother, Athanasia, and my sister, Anastasia, and to my sweetie, Alpay: You are a constant source of encouragement, generosity, love, and inspiration. You are gold to me.

Acknowledgements

Nothing of importance is accomplished alone in this world.

I would like to thank my supervisor Dr. Harvey Kaplan of the Department of Chemistry, University of Ottawa, for making sure I stayed on track with continual guidance. Dr. Kaplan has sparked an interest in me for scientific problem solving, which I can only hope to master as well as he with time and experience in my field of interest. He has afforded me the opportunity to understand that to be successful means to be independently minded, well informed, and to have the ability to utilize, when necessary, the strengths of others to form a solid team effort to solve a research problem. By the end of the entire graduate school process, it is clear that one must draw within oneself and develop a personal philosophy of how to conduct oneself throughout life for the mere sake of survival—everything in moderation should deter disease. I appreciate the fact that Dr. Kaplan has offered me the benefit of the doubt during my trials and tribulations throughout my internship, and have realized with humble gratitude and the greatest respect, that when he was hard on me it was for my benefit and development. I feel proud and privileged to have had such a conscientious supervisor.

I would like to thank my co-supervisor, Dr. Graham Burton, for giving me the opportunity to work at the Steacie Institute for Molecular Sciences, National Research Council of Canada on an interesting and multidisciplinary project, from which a part of this thesis has been comprised. Dr. Burton nurtured my interests with optimism, enthusiasm, respect and encouragement and afforded me the opportunity to experience my first professional scientific work environment. During the course the cholesterol esterase project, there were three key scientists who played a significant role in my progress, namely Dr. Prabhat Arya, Dr. Andrew Moore, and Mrs. Ula Wronska, and I would like to extend my gratitude for their selfless teachings and their support over the years.

I would like to thank our collaborators on a portion of this thesis, namely Mr. Bruce Black, Dr. Brian Dawson, and Dr. Mary-Alice Hefford of the Biologics and Genetic Therapies Directorate, Health Canada, for helpful discussions and guidance. A special thank-you is extended to Dr. Brian Dawson and to Mr. Bruce Black for always being accommodating with regard to NMR time, data analysis and processing. Moreover, I wish to especially thank Mr. Bruce Black for being a good friend during some of the difficult periods.

I have been fortunate to work with extremely kind, generous, supportive, encouraging, hard-working and inspirational people, who have provided me with memories that I will always reflect upon fondly. I would like to sincerely thank my labmates and fellow researchers, Dr. Alpay Taralp (scientific genius), Dr. François Clairmont, Mr. Barry Lorbetskie, Mr. Nicolas Stewart (bright young scientist), Dr. Stephen Lang (NMR expert; scientific genius), Mr. Van Thong Pham (computer genius; extremely generous with his time), Miss Brigitte Simons (newest member of our group), Dr. Illimar Altosaar (worldly), Dr. J. P. (Hans) Schernthaner (always generous

with his time), Dr. Gabriel Alvarado (industrious), Dr. Ravinder Sardana, Mrs. Mary King (kindhearted spirit), Mrs. Mitra Panahi, Mr. Zaman Alli, Dr. Xiongying Cheng, Mr. Maximilian Callaghan (M.D. to be), Dr. Barry Flinn.

I wish to especially thank my soul mate, mentor, loyal and best friend, and sweetie, Dr. Alpay Taralp, for being a constant source of inspiration and encouragement. After all these years, you're still the one. I'll treasure the memories we've made together over the past decade always. With regard to the work at hand, I cannot thank you enough for your endless help and support, and only hope that I've made you proud. The unique humour, kind-heartedness, optimism, enthusiasm, dedication, extreme generosity, confidence, and pride of this young scientific genius will brighten up any work environment, as he has brightened ours.

I wish to thank my truest and dearest kind friend, Ms. Panagiota Karaiskos, who is always just a phone call away when I need to kvetch. Friendship like ours is truly rare. I would like to thank my dear friend, Dr. Christine Allen, who has been understanding of my trials and tribulations, as we share the experience of graduate school. A special thank-you is in order for my kind generous cousin, Louis (Loney) Levetsovitis, who has always been willing to help with computer problems. Thanks for the freshly brewed premium hazelnut coffee from your personal stash while I was working late processing data at Health Canada. Your unique humour and the long talks about life were always a stress reliever. Of course, thanks for the countless rides home. I wish to thank my good friend Mr. George Gharib for his help and support throughout my studies.

A special thank-you is extended to members of the Faculty of Science, in particular from the Department of Chemistry, University of Ottawa (unless otherwise noted), who have encouraged, taught well and with enthusiasm, and/or went out of their way to help during the course of my studies, namely, Dr. Glenn Facey, Dr. Rene Roy, Dr. Alex Fallis, Dr. Tony Durst, Dr. Alain St. Amant, Dr. Hollebone (Carleton University), Dr. Stanley Tsai (Carleton University), and Dr. Denis Rancourt (Department of Physics, University of Ottawa).

I would like to sincerely thank Mr. Guner Taralp, Mme. Berte (Birdie) Proux, Mrs. Aysin Taralp, and Mr. Tarkan Taralp, who have always made me feel welcome in their homes, have always treated me like family, and have extended their support for my studies over the years.

To my parents, Theodore and Athanasia Vakos, and to my sister, Anastasia T. Vakos, with sincere gratitude and the deepest respect, thank-you from the bottom of my heart for all of your sacrifices, for the opportunity to study my inherent interests, for the desire to strive to be the best one can be, for your nurturing, generosity, and belief in education.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

$A_{AC}1$	acid-catalyzed esterification or hydrolysis via unimolecular acyl cleavage
$A_{AC}2$	acid-catalyzed esterification or hydrolysis via bimolecular acyl cleavage
$A_{AL}1$	acid-catalyzed esterification or hydrolysis via unimolecular alkyl cleavage
AOT	dioctyl sodium sulfosuccinate
APEE	*N*-acetyl-L-phenylalanine ethyl ester
ASEE	*N*-acetyl-L-serine ethyl ester
α-TOAc	alpha-tocopheryl acetate
α-TODec	alpha-tocopheryl decanoate
α-TOH	alpha-tocopherol
α-TOPent	alpha-tocopheryl pentanoate
a_w	thermodynamic activity of water: The concentration of water is not an adequate measure of its effect in nonaqueous media, as the influence of water depends not only on the amount of water, but also on the interactions between water and other components in the same phase. It is better to consider the thermodynamic activity of water (a_w). Consider a closed flask with an organic solvent containing dissolved water, some enzyme powder suspended in the solvent, and a vapour space. Water will distribute itself throughout the system until equilibrium is reached. The a_w will be the same in each phase—hydrated enzyme, solvent, and headspace—but the concentrations of water will vary. Thus, equilibrium effects such as enzyme hydration and the balance between hydrolysis and synthesis reactions depend on a_w.
B	buffered sample
BAL	bile acid-stimulated lipase
BCE	bovine cholesterol esterase
Bt	*Bacillus thuringiensis*
Bt-T_1	*Bacillus thuringiensis*-toxin 1
CD	circular dichroism
CE	cholesterol esterase
CGRE	cholesteryl-3-glutaric acid resorufin ester

Ci	Curie
CH	cholate
cPCE	crude porcine cholesterol esterase
cpm	counts per minute
CP/MAS	cross polarization/magic angle spinning
CTAB	cetyltrimethylammonium bromide
DCC	dicyclohexylcarbodiimide
DCU	dicyclohexyl urea
DEPT	distortionless enhancement by polarization transfer
DEPT-135	DEPT conducted at a 135° pulse angle
DFP	diisopropylfluorophosphate
DMAP	dimethylaminopyridine
l-DMPC	*l*-dimyristoyl phosphatidylcholine
D_2O	deuterium oxide
dpm	disintegrations per minute
$\Delta\nu_{1/2}$	linewidth at half-height
eq	electric field gradient
eQ	nuclear quadrupole moment
eqeQ/h	nuclear quadrupolar coupling constant in hertz
ε	extinction coefficient or molar absorptivity
EDTA	ethylenediaminetetraacetic acid
η	deviation of the electric field gradient from axial symmetry = (qxx – qyy)/qzz, where qxx, qyy and qzz are the principal components of the electric field gradient tensor

EPR	electron paramagnetic resonance
ESR	electron spin resonance
E-S	enzyme-substrate complex
FT-IR	Fourier transform-infrared
FT NMR	Fourier transform nuclear magnetic resonance
γ	activity coefficient
GC	gas chromatography
GC-MS	gas chromatography-mass spectrometry
HMED·2HI	hexamethylethylenediamine·2hydriodic acid
HPLC	high-performance liquid chromatography
Hz	hertz
I	nuclear spin quantum number
IR	infrared
$J(\omega)$	spectral density function
kDa	kilodalton
K_m	Michaelis-Menten constant
K_{ma}	aqueous-equivalent K_m (Michaelis-Menten constant)
K_{mo}	Michaelis-Menten constant in organic solvent; = K_{ma} x P, where P is the partition coefficient = $[S]_o/[S]_a$ (a=aqueous solution, o=organic solvent)
K_I	inhibition equilibrium constant
k_{cat}	catalytic turnover
k_{cat}/K_m	specificity constant
$k(\tau)$	auto-correlation function
λ	wavelength

LpH	pH of lyophilization
m_s	electron spin moment
MeI	iodomethane
MHz	megahertz
MS	mass spectrometry
MWCO	molecular weight cut-off
NB	unbuffered sample
NMR	nuclear magnetic resonance
NQCC	nuclear quadrupole coupling constant in hertz
PCE	porcine cholesterol esterase
*pf*PCE	protease-free porcine cholesterol esterase
pD	pD = pH meter reading + 0.4
pH	pH = -log[H^+]
PMSF	phenylmethanesulfonylfluoride
p-NPB	*para*-nitrophenylbutyrate
PNP	*para*-nitrophenyl
ppm	parts per million
SDM	site-directed mutagenesis
SDS	sodium dodecylsulfate
SDS-PAGE	sodium dodecylsulfate-polyacrylamide gel electrophoresis
S_N2	bimolecular nucleophilic substitution reaction
STI	soybean trypsin inhibitor
T_1	longitudinal or nuclear spin-lattice relaxation time

T_2	transverse or nuclear spin-spin relaxation time
τ_c	effective correlation time for reorientation at the NMR active atom
TC	taurocholate
T_2O	^3H_2O, tritiated water
THF	tetrahydrofuran
TLC	thin-layer chromatography
TMA+	tetramethylammonium cation
TMACl	tetramethylammonium chloride
TRIS	tris(hydroxymethyl)aminomethane
UV	ultraviolet
W_0	ratio of water to surfactant molecules in a reverse micelle
ω_0	Larmor frequency for an NMR active nucleus in hertz

Chapter 1

Approaches to the Elucidation of Protein Structure and Function

Research is ongoing to understand the properties of proteins and the relationship between structure and function for the intent of promoting and developing applied protein and enzyme technologies (Means and Feeney, 1971, 1990; Lundblad, 1995). The field of protein chemistry uses physical and chemical techniques to explore protein systems (Means and Feeney, 1971; Darbre, 1986; Lundblad, 1995; Glasel and Deutscher, 1995). Protein analyses are often a composite of these two types of studies, as structure determines function and function depends on structure, and such investigations provide insight into both static (e.g., X-ray crystallographic analysis provides three-dimensional protein structural information) and dynamic components (e.g., kinetic analysis provides catalytic functional information; a bioassay could provide functional information without necessarily providing structural information) of the protein system. Experiments central to protein chemistry probe the interdependence of structure and function in proteins and are often referred to as structure-function studies (Boyer, 1970). The observation that microenvironment can affect intrinsic functional group properties is the driving rationale behind structure-function investigations of proteins (Fersht, 1985; Creighton, 1993; Young and Kaplan, 1989; March, 1988).

1. Established Methods to Study Proteins

Protein chemists have a multitude of techniques and strategies available to be employed collectively as tools to elucidate structure-function relationships, and many volumes of the journal *Methods in Enzymology* have been devoted to this area of protein

research. Physical methods, which are usually spectroscopic measurements (Wyckoff *et al.*, 1985; Hirs and Timasheff, 1986a,b; James and Oppenheimer, 1994; Sauer, 1995; Glasel and Deutscher, 1995), provide structural information, while chemical methods, which are usually covalent chemical modification studies (Means and Feeney, 1971; Lundblad and Noyes, 1984; Lundblad, 1995; Imoto and Yamada, 1989) probe structure and function.

1.1 Physical Methods for Elucidating Protein Structure

An understanding of biological processes involving proteins at the molecular level requires knowledge of the three-dimensional structure of the protein molecule, its dynamic properties, and the intermolecular interactions in which the protein molecule participates. The two methods presently available for determining three-dimensional structures of proteins at the level of atomic resolution are X-ray diffraction and nuclear magnetic resonance (NMR) spectroscopy. Although X-ray crystallography can provide precise structural information, many biological molecules of interest do not yield crystals suitable for X-ray work, or there may be significant differences between structures in a crystal and in solution. The use of NMR spectroscopy does not require crystals, and molecules can be examined in solution under near-physiological conditions where effects of solvent composition, pH, temperature, and other factors on structure can be probed without perturbation of the native conformation (Wüthrich, 1986). Dynamic processes occurring over a wide range of time scales (picoseconds to seconds) can be studied by NMR spectroscopic methods. In the past decade or so NMR spectroscopic methods have been developed to the level at which secondary and tertiary structures of proteins can be determined; two-dimensional (2D) and multidimensional Fourier transform NMR

(FTNMR) spectroscopy along with uniform isotope labeling with ^{13}C and/or ^{15}N are making macromolecules amenable to study (Wüthrich, 1986; Clore and Gronenborn, 1991a,b; Roberts, 1993; Evans, 1995). Information derived from NMR parameters, such as chemical shifts, lineshapes, scalar spin-spin coupling constants, and rates of spin relaxation, are proving useful for deducing the structural features of biological macromolecules. Moreover, developments in instrumentation and methodology and in computer modeling of molecular structure have provided the tools needed for the determination of protein structures by NMR spectroscopy (Glaser and Deutsher, 1995).

It remains today that the major difficulty in studying biopolymer structure by NMR spectroscopy is resolution, which limits the size of the protein that can be studied (i.e., to about 25 kDa). Broad lines due to long correlation times (Harris, 1983) also make the interpretation of spectra difficult. A practical limitation of the use of NMR spectroscopic techniques for protein structure analysis is the fact that the latter are inherently insensitive and high protein concentrations are required, which may be impractical due to solubility considerations or availability of protein.

X-ray and neutron diffraction are unique molecular biophysical methods in that they can be used to develop a structural model that positions tens of thousands of atoms in large macromolecules to within tenths of angstroms. In X-ray diffraction, the structural models are derived from the interference patterns of radiation scattered from crystals of the material (Drago, 1992; Glaser and Deutcher, 1995). The presence of solvent in the crystals of proteins minimizes the conformational perturbations produced by crystallization, because intramolecular interactions vastly outnumber intermolecular interactions (Glaser and Deutscher, 1995).

Other methods that are used to probe the structures of biological molecules are either more limited in the range of problems to which they can be applied or are of lower resolution. Fluorescence energy transfer yields distances between fluorescence donors and acceptors that are correct to within a few angstroms when these groups are separated by less than 80 Å; scanning tunneling microscopy and atomic force microscopy, which measure the force between a probe and individual atoms in a structure, yield structures at resolutions of several angstroms. However, none of these methods, including NMR spectroscopy, has the generality or resolution of X-ray and neutron diffraction, which can be used to determine structures of any size up to a resolution of 1.2 Å. Although X-ray crystallography is the major structural tool, it requires substantial amounts of protein, labour and skill in order to obtain X-ray quality crystals. It is for this reasons that X-ray data of many proteins is not as yet available.

The classical light spectroscopies involve the interactions of electromagnetic radiation with atoms and molecules via absorption, emission, and scattering of photons that provide highly sensitive probes of atomic and molecular structures (Glaser and Deutscher, 1995; Campbell and Dwek, 1984). The classification of the spectroscopic method depends on the wavelength (λ) of the absorbed or emitted radiation. Optical and vibrational spectroscopies include infrared (IR), ultraviolet (UV), and Raman (involves the inelastic scattering of radiation). Subclasses of some spectroscopic methods exist and are labeled according to the particular mechanism of interaction of photons with matter. For example, UV spectroscopy includes the specific techniques of fluorescence emission and circular dichroism, and Raman spectroscopy is often subclassified into off-resonance

and resonance Raman processes. Herzberg (1945, 1950, 1966) has provided a detailed account of the theory and practice of optical and vibrational spectroscopy.

Optical and vibrational spectroscopic methods do not yield directly the three-dimensional structures of proteins. However, the local structures of functionally important groups within a protein molecule, and changes in these local structures that may relate to biological functions, are conveniently probed by optical and vibrational spectroscopy. IR, Raman, circular dichroism and optical rotary dispersion spectroscopies have been useful in elucidating the type and composition of secondary structural elements of protein (Choma *et al.*, 1990), and UV spectroscopy has been used in applications ranging from simple protein quantification, to probing the folding pathway of a protein (Serrano *et al.*, 1992). Experimentally, spectroscopic methods are highly versatile, being applicable to a number of different sample morphologies and over wide ranges of sample temperatures, concentration, and solvent environment. These approaches are also well suited to probing an extensive range of biodynamical processes, with time scales ranging from 10^{-13} sec to days.

Electron spin (or paramagnetic) resonance spectroscopy is a branch of spectroscopy in which radiation of microwave frequency is absorbed by molecules, ions, or atoms possessing electrons with unpaired spins. There are some similarities between NMR and EPR (or ESR) spectroscopy (Drago, 1992); in NMR spectroscopy, the two different energy states (when $I = ½$) arise from the alignment of the nuclear magnetic moments relative to the applied magnetic field, and a transition between them occurs upon the application of radio-frequency radiation. In EPR, different energy states arise from the interaction of the unpaired electron spin moment ($m_s = \pm ½$) with the magnetic

field, the so-called electronic Zeeman effect. EPR measurements are highly sensitive and are of great practical utility for probing local protein structure. Usually a protein is spin-labeled (e.g., with nitroxide) and the motion of the spin-label probes its local microenvironment as a function of solvent properties (e.g., dielectric constant).

Other physical methods include microcalorimetry (Hirs and Timasheff, 1986b), electrophoresis (Leggett Bailey, 1962; Hirs and Timasheff, 1972; Hames and Rickwood, 1990; Karger and Hancock, 1996a,b), hydrodynamic techniques (Hirs and Timasheff, 1985; Glasel and Deutscher, 1995), numerous chromatographies (Hirs, 1967; Hirs and Timasheff, 1972; Darbre, 1986), Mössbauer spectroscopy (Drago, 1992; Glasel and Deutscher, 1995), mass spectrometries (Darbre, 1986; McCloskey, 1990; Biemann, 1992; Drago, 1992; Glasel and Deutscher, 1995). As many techniques are available, the information provided by any one method requires complementation by other techniques.

1.2 Chemical Methods for Elucidating Protein Structure-Function Relationships

There are several chemical modification reagents, enzymes and strategies available to the protein chemist (Leggett Bailey, 1962; Hirs, 1967; Boyer, 1970, 1971; Means and Feeney, 1971; Hirs and Timasheff, 1972; Duggleby and Kaplan, 1975a, Glazer *et al.*, 1976; Lundblad and Noyes, 1984; Darbre, 1986; Imoto and Yamada, 1989; Young and Kaplan, 1989). As with physical methods, individual chemical methods provide only partial characterization in an overall strategy. Unlike physical methods, chemical modification techniques are intrusive, in that they modify protein functional groups and possibly affect interactions that maintain protein structure. Depending on the focus of the investigation, steps may or may not be taken to ensure structural integrity following modification.

Structural integrity is not a priority in a primary structure study, where the N-terminal residue is chemically modified at the α-amino group with fluoro-2,4-dinitrobenzene or dansyl chloride to form acid-stable linkages (Creighton, 1993). The derivatized protein is then hydrolyzed to its constituent amino acids, and the amino acid that is labeled on the α-amino group is detected by the yellow colour in the fluorodinitrobenzene procedure or by fluorescence when the dansyl derivative is detected. While the latter chemical modifications are detrimental to protein structure, they do not defeat the purpose of the investigation, which is to elucidate the primary protein structure. Structural integrity must be considered when the chemical modification study is used to identify functionally important amino acid residues in enzymes (or proteins, since structure-function studies of proteins by chemical modification need not be restricted to the study of enzymes) (Cohen in Boyer, 1971; Means and Feeney, 1971; Lundblad and Noyes, 1984; Darbre, 1986; Lundblad, 1995). Another example where the maintenance of protein structure is unnecessary is the amino acid analysis of a protein, which requires that a protein be acid-hydrolyzed to the individual amino acids (Darbre, 1986). The competitive labeling technique (Duggleby and Kaplan, 1975; Young and Kaplan, 1989) is a double isotope procedure that can be used to distinguish the effects of greater or lesser accessibility from those due to differences in basicity and/or nucleophilicity. It precludes the requirement for structural verification because of the exceptional nature of the experimental design, the reaction conditions, and the trace amounts of reagent employed; the experimental design of this technique is distinct among other chemical approaches since a label can be incorporated into a protein, yet its native properties could be probed (e.g., Hasnain *et al.*, 1977 elucidated the individual pK_a values

for every reactive amino group of a particular protein in a 50-protein ribosome complex under native conditions; the inability of some residues to incorporate label indicated that they were either buried or tightly associated to other proteins).

The characterization of protein function is generally reached via a response to a perturbation of structure (e.g., after the protein has been chemically modified or modified via site-directed mutagenesis; Profy and Schimmel, 1988). The incorporated reagent serves as a probe of structure and function. For example, the competitive labeling technique (Duggleby and Kaplan, 1975; Young and Kaplan, 1989) affords the quantification of pK_a values and structure-induced perturbations of functional group reactivity for each derivatizable group in a protein. Thus, chemical modification of proteins provides information that could not be readily obtainable by physical methods or enzyme kinetics alone. The power of chemical methods has been exploited by the strategic approach that utilizes chemical modification in combination with enzyme activity measurements (Boyer, 1971; Means and Feeney, 1971).

The use of enzyme kinetics exploits the fact that enzymes have an easily measured activity, which in itself served as a probe of structural integrity and the functioning of the various residues (Creighton, 1993). Classically, these amino acid residues contain the side-chain functional groups that participate in bond-making or bond-breaking reactions, and since catalysis and binding are not necessarily separable, amino acids involved in substrate, transition state, or product binding may also be included. Enzyme kinetics has provided information that would otherwise have not been possible for the interactions of an enzyme with substrates, inhibitors and affinity labels. Kinetic studies have been part of experimental strategies that made use of

isotopes and different dielectric media (Purich, 1980; Cleland, 1982; Wangikar *et al.*, 1993, 1995). In this manner, enzyme kinetics could be used to investigate functional groups that are essential to binding and/or catalysis, to probe transition state energetics and reaction pathways, and to differentiate between different classes of ionizable groups involved in catalysis.

When covalent modification of a particular residue results in the loss of enzymatic activity, it is generally assumed that its side-chain functional group is "essential" for catalysis. The rate of modification provides additional information about the reactivity of the important amino acid and it is possible that the residue could be located in the primary structure of the protein by peptide mapping. However, this approach is subject to several well-known limitations inherent to chemical modification, which can complicate the interpretation of functional results (Profy and Schimmel, 1988). The most important of these are as follows.

(1) The first is the limited number of structural changes that can be produced. Almost invariably, labeling results in an increase in side-chain size. Some enzymes have been inactivated by bulky reagents, but not by small ones.

(2) The second is the incomplete specificity of modification reagents. Chemical modification reagents rarely label a unique residue. The most common reagents react predominantly with one type of side-chain, but complete specificity for a given type is rare. Even when the reagent is highly side-chain specific, several side-chains of the corresponding type may be modified. If activity is lost in this case, it may be postulated that a residue of a particular type is essential, but the actual residue may be unidentifiable. Sometimes, the structure of the native enzyme enhances the reactivity of

active site residues so that they are labeled much more quickly than other residues of the same type [e.g., the selective phosphorylation of the active site serine residue in chymotrypsin with diisopropylfluorophosphate (Jansen *et al.*, 1949)]. Affinity labels may react with residues at the binding site, or may be activated *in situ* (e.g., by photolysis) to label nearby side-chains. The mechanism-based inhibitors, which are a class of affinity labels, are activated for covalent modification by undergoing part of the normal catalytic reaction. It is also possible to label an active site residue specifically by trapping a covalent enzyme-substrate complex through chemical treatment. Despite these methods for improving specificity, most modifications lead to the labeling of more than one amino acid in an enzyme. It is sometimes possible to correlate the rate of inactivation to the rate of appearance of a peptide that contains the labeled residue, but when several residues react at the same rate, the important one cannot be identified.

(3) The third is the possibility that modification has not been quantitative; functional conclusions are limited by the *extent* of modification. If a particular side-chain has not been quantitatively derivatized, conclusions about its role in catalysis are weakened by the possibility of contamination by unreacted enzyme. A large effort must be expended to show that labeling is quantitative.

(4) The fourth is the possibility that modification results in the disruption of the spatial arrangement of active site atoms following a global structural change rather than a local perturbation due to the derivatization of a catalytically essential side-chain. Such a structural change could occur upon modification of residues near to or distinct from the active site, and could result from, for example, the loss of a hydrogen bond or the introduction of steric bulk. It is also possible that inactivation of an enzyme could result

from the introduction of steric bulk at a nonessential residue near the active site. Although this modification may not alter the orientation of active site groups, it may physically block substrate binding. This problem is exacerbated by the fact that most modifications lead to an increase in side-chain size.

As a result of the aforementioned limitations, conclusions based on chemical modification evidence have at best been tentative, and it has been desirable to support the identification of important residues with corroborating data. The advent of site-directed mutagenesis (SDM) has provided the protein chemist with the opportunity to make defined amino acid substitutions at specific positions in enzymes. Preselected codon changes can be introduced into the coding region of the cloned, sequenced gene that codes for the enzyme of interest such that the expressed protein contains the corresponding amino acid replacement (the mutant gene is expressed in the absence of a wild-type or homologous copy; synthetic oligonucleotides are required). This alternative method of perturbing protein structure is useful for enzyme structure-function studies because it overcomes many of the limitations associated with chemical modification. In particular, the amino acid substitutions are limited to defined positions (specific and quantitative), and the differences between the native and mutant protein are subtler; this overcomes the problem of side reactions that plague chemical modification experiments. However, in order to perform a SDM experiment, it is necessary to identify regions of the enzyme where the effects of amino acid changes will be interesting and interpretable. In cases where a three-dimensional structure of a protein is unavailable, there are clear advantages to combining the techniques of chemical modification and SDM for the study of enzyme function. Classical chemical modification data can play a necessary role in

locating amino acid residues that may participate in catalysis. Then, on the basis of this information, SDM can be used to construct mutant proteins that contain amino acid substitutions at these positions. The functional importance of the altered residues can then be assessed in a relatively unambiguous way by comparing the catalytic properties (or some other function) of the mutants to those of the native enzyme (or protein).

Another major advantage of SDM is that a wider variety of chemical changes can be made. In chemical modification studies, although a large number of chemical modification reagents have been designed, only a few types of amino acid residues are susceptible to labeling. In SDM studies, particularly significant is that smaller side-chains can be introduced at desired positions. Although smaller residues may also affect structural integrity, the possibility of steric occlusion of the active site is eliminated. SDM can be used to generate subtle changes in protein structure. By appropriate selection of amino acid replacements, the effect of changing only one or a few atoms can be studied. For example, Tyr→Phe and Ser→Ala changes can be used to determine the role of specific hydroxyl groups, and Glu→Gln and Asp→Asn changes can probe the importance of the negatively charged carboxylate groups. The possible replacements are limited to the other 19 amino acids. However, it is possible to introduce a residue such as cysteine that can be derivatized by several chemical modification reagents and the properties of these labeled mutants may be revealing.

As with chemical modification, in a SDM study it is often unclear whether a change in enzymatic activity results from the replacement of an essential side-chain functional group or from a global structural change; however, the latter is less likely in SDM studies because the modifications are more subtle and specific, as evidenced by

some X-ray crystallographic studies; nevertheless, the possibility of global structural change is difficult to rule out without some structural information.

SDM provides a useful test of the validity of conclusions based on chemical modification experiments. It is possible to replace a number of putative "essential" residues without a substantial activity loss (Profy and Schimmel, 1988). This demonstrates the power of SDM as a structure-function probe. It is also expected that replacement of some "essential" amino acids would cause a dramatic loss in enzymatic activity, thereby confirming their importance. An important finding is that catalysis can still be observed when residues evidently important for catalysis are replaced by amino acids of similar structure by SDM. This shows that instead of thinking in terms of essential amino acid residues, it is more appropriate to consider essential functionality. Moreover, the observation of residual (low, but measurable) activity in mutant proteins may provide a useful new tool. If replacement of a residue suspected in having a functional role by one of similar structure produces a mutant with residual activity, while its replacement by other amino acids results in inactive mutants, then the residue may well be involved in catalysis; but, it cannot be ruled out that the residue may be involved in a structural role, for example, in the maintenance of secondary structure.

The classic protein titration assay (Hirs, 1967; Means and Feeney, 1971) is a useful method to obtain structure-function information. The nature and number of ionizable groups in a protein can be determined from the thermodynamic titration profiles; pI values could also be estimated. Titration hysteresis gives additional insight into whether groups are surface exposed or buried.

2. Use of Low Water Systems to Study Proteins

Since enzymes have evolved in a water milieu, it is not surprising that most of the current knowledge of enzymes has come from conventional experiments conducted in aqueous solutions. It is clear that water is a fundamental requirement for enzyme function and for the formation of the three-dimensional structure of proteins (Rupley *et al.*, 1983). However, it is assumed that conventional studies limit the information that can be obtained on the relationship between enzymes and water, since water is usually the most abundant component of the experimental system. Consequently, it became necessary to design novel experimental systems that use water as a variable parameter to provide further information on enzyme function, structure, and dynamics. Presently, there are two experimental systems that are widely employed for the study of the interaction between water and enzymes. One involves the dispersion of enzymes in nonpolar solvents (Klibanov, 1989), and the other the entrapment of enzymes in reverse micelles (Luisi *et al.*, 1988). In these two nonconventional systems the amount of water in which enzymes are placed is much lower than in the usual biochemical systems. Several reviews have dealt with the characteristics of these systems and the behaviour of enzymes in such conditions (Luisi *et al.*, 1988; Martinek *et al.*, 1989; Pileni, 1989; Gupta, 1992; Dordick, 1989, 1992; Gómez-Puyou *et al.*, 1992; Westcott and Klibanov, 1994; Nicot and Waks, 1996; Bru *et al.*, 1995; Tuena de Gómez-Puyou and Gómez-Puyou, 1998). These studies have revealed properties of enzymes that are not easily appreciated from experiments conducted in aqueous solutions.

In the low water systems, it has been possible to probe the relation between solvent and enzyme kinetics, as well as some of the factors that affect enzyme thermostability and catalysis. Furthermore, the studies show that low water environments

can be used to stabilize conformers that exhibit unsuspected catalytic properties, as well as intermediates of enzyme function and formation that in aqueous media have relatively short life-times. When catalysis by an enzyme placed in low water systems is observed, the obvious question that comes to mind is whether in such conditions the enzyme has, or does not have, a structure similar to that in solution. In either case, for full understanding of enzymes it is imperative to determine and compare their structures in the two conditions. The study of the structure of enzymes in these unnatural conditions is an active field of research.

A description of the function and structure of enzymes dispersed in organic solvents is presented herein; where relevant, that of enzymes in reverse micelles is included. This background information will lead to an overview of the innovative studies of Taralp and Kaplan (1997) on the *in vacuo* chemical modification approach, which is a third nonconventional low water system by which researchers can study enzymes, and proteins in general, for the elucidation of structure-function relationships. Applications and development of this third approach are emphasized in the present body of research in Chapters 2-5. A review of the background information on the nonaqueous experiments with enzymes dispersed in organic solvents is necessary to justify the *in vacuo* approach (i.e., lyophilized enzymes chemically reacted at high temperatures). An outline of the information that can be observed from studies on enzymes entrapped in reverse micelles is more appropriately presented in Chapter 6, where a separate study on cholesterol esterase is described. A review of the experiments with enzymes entrapped in reverse micelles is helpful to appreciate the direction in which the micellar enzymology experiments conducted with cholesterol esterase can take for future research. The low

water reverse micellar system may be employed as a tool to inquire into parameters that in biological systems are difficult to determine. The chapter is concluded with a description of each of the projects of the present research.

2.1 Characteristics of Enzymes Dispersed in Organic Solvents

The general protocol for measuring enzyme activity in organic solvents involves the lyophilization of the enzyme from aqueous solution of a particular pH (for optimal activity) (Zaks and Klibanov, 1988a). The enzyme powder is suspended in an organic solvent that has a known amount of water; the substrate is added subsequently. Samples are vigorously shaken during the incubation period and aliquots are withdrawn for assay of product at appropriate times. When a water-soluble cofactor is required, it is usually co-lyophilized with the enzyme. The most important finding stemming from the study of enzymes in organic solvents is that water, although in relatively low amounts in such systems, is an essential ingredient for enzyme catalysis. Several convincing reports have demonstrated that enzymes are active with amounts of water that are below monolayer coverage (Zaks and Klibanov, 1988a, 1988b; Dordick, 1992; Adlercreutz, 1991).

Although enzymes have been shown to work in organic solvents, it was necessary to determine whether their bound water molecules were essential for catalysis or whether they merely resist removal by organic solvents. The studies of Zaks and Klibanov (1988a, 1988b) have determined the activities of three unrelated enzymes (1988a; alcohol oxidase, polyphenol oxidase, and alcohol dehydrogenase) and two serine proteases (1988b; subtilisin and α-chymotrypsin) dispersed in different organic solvents and different amounts of water. Activity increased with the amount of water in the system in all solvents, but significantly, less water was required for catalysis with the more

hydrophobic solvents. Enzyme activity depended on the amount of water that was bound to the enzyme, and not on the overall content of water in the system.

Gorman and Dordick (1992) explored the stability of bound radiolabeled water (T_2O) in α-chymotrypsin, subtilisin Carlsberg, and horseradish peroxidase in different organic solvents. Removal of bound T_2O was related to the dielectric, partition coefficient, and capacity of the solvent to solubilize water. The nature of the enzyme was also a factor in the stability of bound T_2O, but the contribution of the solvent was more important. The overall data led to the concept that in organic solvents of low water activity, enzymes work better in nonpolar than in polar solvents because of their low capacity to remove or solubilize the tightly bound water molecules.

Direct evidence for the existence of water molecules that resist removal by organic solvents has come from analysis of the crystal structure of enzymes that have been crystallized in water and subsequently exposed to organic solvents (Fitzpatrick *et al.*, 1993, 1994; Allen *et al.*, 1996; Yennawar *et al.*, 1994, 1995; Schmitke *et al.*, 1997). Water molecules were observed in particular regions of the protein. Solvent molecules were detected to be in close range of protein residues. Thus, enzymes dispersed in organic solvents need some essential water molecules for expression of catalysis; the essential water maintains the protein structure.

2.1.1 Effect of Water and Solvent on Catalysis by Enzymes Dispersed in Organic Solvents and in Reverse Micelles

Enzymes in all low water systems exhibit activities that are much lower than in aqueous media in the very low range of water concentrations. As water is increased, activities also increase. The activity of enzymes dispersed in organic solvents or entrapped in reverse micelles often respond with a different pattern to increasing water

concentrations. The unique features that enzymes exhibit in the two systems provide insight into properties of enzymes that are not easily observed in aqueous solutions. A description of the activity of enzymes entrapped in reverse micelles is presented in Chapter 6 and that for enzymes dispersed in organic solvents is dealt with below.

2.1.2 Characteristics of Catalysis by Enzymes Dispersed in Organic Solvents

Chymotrypsin and xantine oxidase were observed to be active in organic solvents in 1966 and 1967 by Dastoli, Musto, and Price. Some 20 years later, Zaks and Klibanov (1984, 1985) pursued these observations to revive the field of nonaqueous enzymology. They reported that lipase dispersed in organic solvents exhibited catalytic activity at a temperature of 100°C. Many enzymes, differing markedly in structure and catalytic mechanisms, exhibit catalysis under these conditions, which intuitively seemed deleterious for enzyme action and structure. For example, subtilisin, α-chymotrypsin (Zaks and Klibanov, 1988b, Westcott and Klibanov, 1993), lipases (Zaks and Klibanov, 1985), yeast alcohol dehydrogenase, polyphenol oxidase, horse liver alcohol dehydrogenase (Zaks and Klibanov, 1988a), horseradish peroxidase (Ryu and Dordick, 1992), and alkaline proteases (Tawaki and Klibanov, 1992) are active in organic solvents.

2.1.2A. Kinetic Studies

Enzymes dispersed in organic solvents exhibit Michaelis-Menten behaviour in the range of substrate concentration that can be assayed. Several factors strongly influence catalysis (Ryu and Dordick, 1989, 1992). Ryu and Dordick (1992) investigated linear free energy relationships between horseradish peroxidase activity (catalyzing the oxidation of phenols) and the properties of substrates and solvents. They demonstrated that the substantial increase in the K_m as both substrate and solvent hydrophobicities

increased (i.e., the partitioning of substrate is reduced, thereby requiring a higher concentration of phenols to saturate the enzyme) is a result of the reduced binding efficiency of phenolic substrates due to the enhanced stabilization of substrates in the ground state, rather than the destabilization of the E-S complex. Their work showed that catalytic efficiency depends on the physicochemical characteristics of both substrate (hydrophobicity and electronic properties) and solvent (hydrophobicity and polarity).

Wangikar *et al.* (1993) also have demonstrated that the reactivity of an enzyme is very much dependent on the polarity of the solvent, substrate, and enzyme active site. Since subtilisin-catalyzed transesterifications in organic media follow the acyl-enzyme mechanism involving a charged and polar tetrahedral transition state, it is not surprising that a polar environment stabilizes the transition state. Indeed, the transition state of subtilisin BPN' is intrinsically more stable in acetone (polar) than in hexane (nonpolar) (the difference in intrinsic free energies of transition state stabilization, $\Delta\Delta G_{ES}^{\ddagger}$, between hexane and acetone is 1.75 kcal/mol). A polar active site mutation (Gly166 to Asn or G166N) can compensate for the transition state destabilization (of the polar substrate *N*-acetyl-L-Ser ethyl ester) upon going from acetone to hexane, and thus stabilize the intrinsic transition state of the enzyme in hexane by 1.22 kcal/mol as compared to acetone. The polar substrate partitioned favourably from the nonpolar solvent to the polar active site. The ground state was destabilized in hexane but the polar transition state was stabilized in hexane relative to acetone (compared with the wild type enzyme) by the polar mutant enzyme. Thus, the mutation, while deleterious in aqueous solutions, resulted in improved catalysis in hydrophobic organic solvents. The polar mutant also stabilizes the transition state of a polar substrate (*N*-acetyl-L-Ser ethyl ester; ASEE) as

compared to that of a nonpolar substrate (*N*-acetyl-L-Phe ethyl ester; APEE) in a range of different organic solvents. The transition state stabilization due to the G166N mutation must be somewhat local to the S_1 acyl-binding site; perhaps an enhanced interaction such as H-bonding between the amide of Asn and the hydroxyl function of ASEE takes place and stabilizes the enzymatic transition state; such a polar interaction is not possible with the nonpolar substrate (APEE) with the G166N mutant. The transition state stabilization correlates well with solvent polarity rather than solvent hydrophobicity, which suggests that the solvent effect is primarily electrostatic, as expected for a charged transition state. In this way, subtilisin BPN' is significantly activated (up to 178-fold) by a polar active site mutant in going from acetone to hexane.

An intriguing question arising from the results of Wangikar *et al.* (1993) was whether water affects the kinetics of the enzyme at the catalytic site. Addition of water can also improve the degree of enzyme catalysis (i.e., transition state stabilization, decrease in $\Delta G^{\ddagger}$; increase in k_{cat} or catalytic turnover) by increasing the acylation rate and enzyme-substrate binding (i.e., decreasing the intrinsic K_m value). Importantly, in dry tetrahydrofuran, acylation is rate-limiting, whereas in the same solvent with 2% water, deacylation became rate-limiting. Significantly, Chang and Shiao (1994) reported that the entrapment of alkaline phosphatase in reverse micelles causes a shift of the rate-limiting step.

Xu *et al.* (1994) continued the investigation on transition state stabilization of subtilisins in organic media, and again confirmed that polarity is an important factor in stabilizing the charged transition state, and that the rate of an enzymatic reaction proceeding through a charged transition state can be enhanced by increasing the active

site polarity (e.g., by the addition of water). Their kinetic results for subtilisin BPN'-catalyzed hydrolysis of *N*-acetyl-L-phenylalanine ethyl ester in aqueous buffer, and for the transesterification reaction with 1-propanol in tetrahydrofuran (THF) revealed the important fact that enzymes do not work as well in organic solvents as they do in water (k_{cat}/K_m in aqueous buffer 2840 $M^{-1}s^{-1}$; k_{cat}/K_m in THF 0.081 $M^{-1}s^{-1}$). At least part (Klibanov, 1997) of this difference represents the loss of free energy of stabilization in the transition state between the aqueous and THF reactions [$\Delta G^{\ddagger}$ increases and k_{cat} decreases as a result of an increase in K_m; observed effect due to differences in solvation of the substrate in THF (stabilization of apolar substrate in the ground state) compared to the situation in water]. Thus, dehydration of subtilisin may particularly affect the active site region and destabilize the charged transition state of the enzyme-catalyzed transesterification reaction. The lost free energy of transition state stabilization may be partially restored by increasing the hydration of the enzyme. Indeed, a tenfold increase in the catalytic efficiency of subtilisin BPN' was observed with the addition of the 5 μL/mL water to the transesterification reaction of *N*-acetyl-L-phenylalanine ethyl ester with 1-propanol in THF. Since the acylation step proceeds through a charged tetrahedral transition state, an increase in active site polarity should stabilize the transition state and accelerate the rate of reaction. Site-directed mutagenesis studies showed that a polar mutant (G166N) also stabilized the transition state for catalysis in THF relative to catalysis in aqueous buffer, as expected for a charged transition state in nonaqueous media. The greater catalytic efficiency (i.e., transition state stabilization) of G166N in dry THF relative to the wild-type subtilisin BPN' stemmed from its lower K_m value (i.e., improved binding of substrate by polar mutant) and to a slight increase in k_{cat} ratios.

The above results demonstrate that enzyme activity is mainly controlled by the partitioning of the substrate from the solvent to the active site, which determines the ground state level of the substrate (Koskinen and Klibanov, 1996; Halling, 1994). Substrate solvation differences in organic solvents mainly affect the total free energy differences for the enzymatic reaction and are reflected by differences in activity coefficients [$(k_{cat}/K_m)_A/(k_{cat}/K_m)_B = \gamma_A/\gamma_B$], as well as in changes in partition coefficients [K_{mo} is proportional to the partition coefficient, $P = [S]_o/[S]_a$; o = organic solvent, a = aqueous solution; $K_{mo} = K_{ma} \cdot P$ and $K_{mo}/P = K_{ma}$, corrected, aqueous-equivalent K_m value; the latter are compared between substrates with the 'true' K_m value measured in aqueous solution to help determine relative reaction rates and to help explain changes and even reversals in enzyme specificity]. There is a general increase in apparent K_m values when enzymes are placed in organic solvents rather than in aqueous solution (Yang and Robb, 1993).

Xu *et al.* (1994) employed two additional independent methods to probe the role of water in enzyme catalysis. The rate of inactivation of subtilisin BPN' with the active site inhibitor, phenylmethanesulfonylfluoride or PMSF (enzyme is sulfonylated by the inhibitor) was examined; as water is added to THF, the rate of PMSF inactivation increases. The physicochemical properties of the microenvironment of the active site as a function of the water content of the solvent was examined by spin labeling (with nitroxide) Ser_{221}, and conducting EPR spectroscopy; the local polarity increased sharply as water is added up to ca. 20 μL/mL, then levels off at higher water concentrations. Both the PMSF inhibition results and the EPR results independently showed similar

trends as water was added to the THF solvent, such that as the active site polarity increases, so does the catalytic competency of the enzyme.

Wangikar *et al.* (1995) pointed out that although the energetics of hydrophobic interactions between substrate and solvent can be determined by solvation calculations, the contribution of hydrophobic interactions between the substrate and the enzyme are difficult to define due to the uncertainties in the hydrophobic state of the active site. To this end, the free energy of activation was determined for various *N*-acetyl-L-amino acid ethyl ester substrates that differed in amino acid side-chain hydrophobicity in three enzymes. Wild-type subtilisin BPN' and two mutants in which Gly-166 at the active site (the S_1 binding cleft) was substituted by either an Ala or a Val were studied. The substitutions make the catalytic site more hydrophobic and exert only small steric influences on k_{cat}/K_m. Thermodynamic analysis of the data in aqueous media showed that the mutations affected enzyme specificity, whereas in hydrophobic solvents the mutations strongly affected the transition state. The data indicated that in the expression of catalysis by enzymes dispersed in organic solvents, there is a fine balance between at least three components: the polarity of the substrate, the solvent, and the catalytic site. The authors stressed that by adjustments in these parameters it is possible to make quantitative predictions on how to control catalytic rates and substrate specificity.

2.1.2B. Enzyme Specificity

Westcott and Klibanov (1994) have reviewed the notable finding that some enzymes dispersed in organic solvents show altered selectivity or specificity toward a given substrate depending on the solvent in which the enzyme is placed. Solvent engineering is the challenge to make quantitative predictions on how to adjust the

behaviour of the enzyme toward a given substrate (Westcott and Klibanov, 1994; Wangikar *et al.*, 1995; Ke *et al.*, 1996; Westcott *et al.*, 1996; Klibanov, 1997). Assuming that solvent does not modify the catalytic site, the experimental data show that solvation energy contributes strongly to the partition of different substrates into the catalytic pockets. The rational manipulation of solvents to alter both the partition of substrates into the catalytic site and the equilibrium constant of the reaction, in conjunction with modification of catalytic sites by site-directed mutagenesis, are powerful tools to manipulate enzyme action.

2.1.3 Water, Enzyme Flexibility, and Catalysis

Independent of the nature of the experimental system, (e.g. enzymes dispersed in organic solvents or entrapped in reverse micelles), all enzymes exhibit an identical characteristic. When water concentration is below monolayer coverage, or at a very low ratio of water to surfactant, W_0 (3 or below; see Chapter 6), in reverse micelles, the activity of enzymes is many times lower than in conventional aqueous media. Paradoxically, it is in these conditions in which research has given fundamental information on basic enzymology and has opened new panoramas in biotechnology.

The low or lack of activity that enzymes exhibit at low water concentrations (or water activities) is not the consequence of irreversible enzyme denaturation, since an increase in water brings about the appearance of catalytic activity. The lack of substrate binding does not account for a large proportion of the activity loss. For example, the number of competent catalytic sites of crystallized subtilisin suspended in acetonitrile was determined from titrations of the active site serine to be one half of that observed in solution, even though the differences in k_{cat}/K_m was on the order of 10^7 times (Schmitke

et al., 1996). Significantly, binding of the coenzyme of alcohol dehydrogenase entrapped in reverse micelles was similar to that observed with the enzyme in solution, however negligible catalytic activity was observed (with a W_0 of 11) (Strambini and Gonnelli, 1988). Similarly, binding of a substrate analogue to triosephosphate isomerase entrapped in reverse micelles with low water content was unimpaired, but the enzyme showed less than 5% of the activity in solution (Garza-Ramos *et al.*, 1994).

2.1.3A. Enzyme Flexibility

Enzymes have diminished flexibility in low water environments, which leads to low catalytic rates (Klibanov, 1989, 1997; Dordick, 1992). In a low water environment, enzymes are relatively rigid structures (kinetically but not thermodynamically trapped in their native-like conformation; Zaks and Klibanov, 1988) that may reveal properties that would not be appreciated when water is in excess. The structural rigidity of enzymes in organic solvents has been indirectly substantiated by enzyme activity assays (e.g., α-chymotrypsin in octane is dependent upon the pH of the aqueous solution from which the enzyme had been lyophilized, with the 'pH dependence' resembling that in water; Zaks and Klibanov, 1988). The pioneering studies of Rupley *et al.* (1983) with lysozyme powders at various levels of hydration showed the strong dependence of protein mobility and enzyme activity on water. Bone (1987) proposed from time domain spectroscopy studies of α-chymotrypsin that water-induced flexibility of proteins is the consequence of a diminution in the interactions between charged and polar residues of the protein molecule caused by dielectric screening. Water was envisioned as a plasticizer or lubricant that facilitates structural flexibility by shielding interactions between polar residues (Bone and Plethig, 1982). Organic solvents lack the ability of water to engage in

multiple H-bonds and act like a molecular lubricant; they have lower dielectric constants compared to water, which leads to stronger electrostatic interactions in the rigid proteins with reduced catalytic activities (Affleck *et al.*, 1992).

Enzymes are very flexible structures in solution (Karplus and McCammon, 1983), and in a catalytic cycle, there must be a strong interplay between the solvent and many protein groups (Reid and Rand, 1997). During catalysis, enzymes undergo conformational changes when the substrate binds to the active center, when it is transformed into product, and when the latter is ejected into the media. The crystal structure of enzymes, in the presence or absence of active site ligands, show that in some enzymes the conformational changes are rather subtle, but quite large in others. Some enzymes have lids or loops that close over the active site after substrate binding that involve movements of several angstroms (lipase, Brzozowski *et al.*, 1991; triosephosphate isomerase, Wierenga *et al.*, 1991; lactate dehydrogenase, Holbrook *et al.*, 1975).

Burke *et al.* (1993) showed that the flipping rate constant of a tyrosine residue of α-lytic protease suspended in organic solvents is 10^3 s^{-1}, whereas in solution it is four orders of magnitude higher; studies were carried out by solid-state NMR spectroscopy. EPR measurements of spin-labeled residues of α-chymotrypsin in solvents of different dielectric constants showed that motion increases as the dielectric constant of the solvent increases (Affleck *et al.*, 1992a); molecular simulation studies revealed that the residues on the outer regions of the protein were more affected. Guinn *et al.* (1991) have shown by EPR spectroscopy of spin-labeled lactate alcohol dehydrogenase that enzymes are more flexible in solvents with high dielectric constant than in those with low dielectric

constant. The degree of enzyme structural rigidity also depends on the pH of lyophilization; the ionization state of the enzyme influences dynamics in organic solvents (Guinn *et al.*, 1991).

Affleck *et al.* (1992b) carried out a parallel study of the transesterification reaction catalyzed by subtilisin Carlsberg and the dynamics of nitroxide spin label bound to the active site serine in tetrahydrofuran. The progressive additions of water produced an increase in activity that was followed by the diminution and nearly complete abolition of catalysis. The increase in catalytic rates coincided with an increase in active site polarity and an enhancement of mobility of the spin label.

Studies of the dynamics of the spin-labeled active site of α-chymotrypsin entrapped in reverse micelles revealed that the label exhibited a lower mobility than when the enzyme was in water (Marzola *et al.*, 1991). Examination of the intrinsic fluorescence of tryptophan of α-chymotrypsin entrapped in reverse micelles in the very low W_0 range, where the enzyme is apparently in an apolar environment and in a frozen state, revealed that, relative to the enzyme in aqueous media, at W_0 of 0.65, the enzyme exhibited a marked blue shift of its intrinsic fluorescence (Dorovska-Taran *et al.*, 1993). As water concentration was increased, activity appeared and the fluorescence spectra shifted to the red. Significantly, the amount of water at which catalytic activity was induced in reverse micelles and in the enzyme suspended in tetrahydrofuran (Affleck *et al.*, 1992b) was in the same range.

Hence, the relationship that exists between protein mobility, activity, and water has been observed in lyophilized enzymes dispersed in organic solvents and enzymes entrapped in reverse micelles.

2.1.3B. Enhancement of Catalytic Activities by Denaturants

Enzymes in a low water environment have hampered activity due to constraints in conformational mobility. Denaturing agents (e.g., guanidine chloride) that promote solvent-protein interactions have been shown to enhance the catalytic rates of enzymes in low water systems to different extents (Garza-Ramos *et al.*, 1992; Fernández-Velasco *et al.*, 1992).

Activation of enzymes dispersed in organic solvents can be achieved by denaturing cosolvents. Dimethyl sulfoxide and formamide produced a dramatic increase in activity of subtilisin, α-chymotrypsin, and lipase to different extents (Almarsson and Klibanov, 1996). The authors concluded that the denaturing cosolvents brought about an increase in the flexibility of the enzyme. Significantly, methanol produced strong activation of chloroplast ATPase entrapped in reverse micelles (Kernen *et al.*, 1997). Thus, cosolvents can be used advantageously to increase catalytic rates in low water environments.

2.1.3C. Factors that Contribute to the Low Activity of Enzymes in a Low Water Environment

The data presented thus far are in agreement with the idea that in low water milieu the conformational freedom of enzymes is restricted. However, they do not prove that enzyme rigidity is the sole cause of the low activity that enzymes express in low water environments. Schmitke *et al.* (1996; also see discussion by Klibanov, 1997) addressed the question by comparing the activity of subtilisin in solution with that of crosslinked-crystallized subtilisin in acetonitrile in the hydrolysis of *N*-acetyl-L-Phe-OEt. They carried out a mechanistic dissection of the causes for the plunge in enzymatic activity in anhydrous solvents compared to the situation in water. The crystal structure of the

crosslinked enzyme is strikingly similar to that of the non-crosslinked enzyme (Fitzpatrick *et al.*, 1993, 1994; Schmitke *et al.*, 1997). Thus, the crosslinked enzyme crystals placed in organic solvents are a good model system to use for dissecting the energy terms of each contributing factor to the decline of enzyme activity in organic solvents compared to in water. The strategy provided a basis for a comparative study amongst solvents without considering structural differences (Schmitke *et al.*, 1997). The difference in k_{cat}/K_m at pH 7.8 of the enzymes in the two conditions was on the order of 10^7 times. A systematic study of the variations of various parameters led to the observation that three factors contribute to the differences in k_{cat}/K_m: (a) a shift of the activity to a more alkaline pH for the enzyme in the organic solvent; (b) an unfavourable effect on the desolvation of the substrate, and in consequence on its partition to the catalytic site; and (c) a diminished conformational freedom of the enzyme. Experiments on the effect of two solvents with different capacity to form H-bonds revealed that the latter barrier for enzyme catalysis was related to the thermodynamic activity of water. It remains to be established if these mechanisms also operate in other enzymes or whether additional factors (Dong *et al.*, 1996) account for the low activity that enzymes express in low water environments.

Research is ongoing to determine ways to increase the rates of enzyme catalysis in organic solvents (Adlercreutz, 1991; Klibanov, 1997). The most popular strategy has been to prevent enzyme denaturation during dehydration. This has been achieved by lyophilizing enzymes in the presence of lyoprotectants such as sugars, poly(ethylene glycol) (Dabulis and Klibanov, 1993), inorganic salts (notably KCl; Khmelnitsky *et al.*, 1994), substrate-resembling ligands (Dabulis and Klibanov, 1993; Russell and Klibanov,

1987), and crown ethers (Broos *et al.*, 1995). An alternative is to preform organic-soluble complexes of enzymes with detergent (or other amphiphile) molecules (Okahata and Ijiro, 1992; Paradkar and Dordick, 1994); enzymes are apparently in native-like conformations in such complexes. These approaches have resulted in increasing enzymatic activities in organic solvents by up to three to four orders of magnitude. As mentioned above, the other effective strategy has been to 'loosen up' enzymes in anhydrous solvents. Enzyme activity has been increased up to two or three orders of magnitude by adding small amounts of water to organic solvents (Zaks and Klibanov, 1988), maximizing the thermodynamic activity of water, a_w, (Bell *et al.*, 1995), and adding water mimics (Zaks and Klibanov, 1988; Kitaguchi and Klibanov, 1989) or denaturing cosolvents (Almarsson and Klibanov, 1996) to anhydrous media.

It is clear that water is instrumental in protein mobility, as well as on other factors that participate in the expression of catalysis. The analysis and proof of this concept has led to the visualization of many novel and unsuspected characteristics of enzymes, as outlined below.

2.1.4 Water and the Relationship between Stability and Catalysis

Solvation is necessary for protein function, but excessive solvation may lead to loss of the native protein structure. This is more notable at relatively high temperatures in which many of the weak bonds that maintain the native structure are destabilized and solvated. It was rationalized that in a low water environment, enzymes should exhibit a high resistance to thermal denaturation, and indeed, the latter proved correct.

2.1.4A. Thermostability

Several enzymes dispersed in organic solvents with low water concentrations have been shown to exhibit thermostability many times higher than in solution: lipases (Zaks and Klibanov, 1984, 1985), terpene cyclases (Wheeler and Croteau, 1986), α-chymotrypsin (Reslow *et al.*, 1987; Mozhaev *et al.*, 1991), α-amylase (Asther and Meunier, 1990), ribonuclease, and lysozyme (Volkin *et al.*, 1991). Thermostability decreased as the water content of the systems was increased, indicating that an inverse relationship exists between enzyme thermostability and water concentration (Zaks and Klibanov, 1984; Ayala *et al.*, 1986).

For enzymes entrapped in reverse micelles, the nature of the amphiphile is paramount to whether the enzyme exhibits thermostability. Enzymes entrapped in reverse micelles formed with phospholipids exhibited remarkable thermostability (Ayala *et al.*, 1986; Garza-Ramos *et al.*, 1989, 1990, 1994). Some enzymes entrapped in reverse micelles formed with synthetic detergents resulted in a lower thermostability than in solution (Luisi *et al.*, 1990; Khmelnitsky *et al.*, 1993; Dorovska-Taran *et al.*, 1993; Chang and Shiao, 1994), however one example exhibited a high thermostability (e.g., catalase entrapped in reverse micelles formed with dioctyl sodium sulfosuccinate or AOT; Escamilla *et al.*, 1992).

2.1.4B. Cleavage of Covalent Bonds in Low Water Systems

Thermal denaturation of enzymes in aqueous media is often accompanied by cleavage of covalent bonds. In a low water environment, enzymes become more resistant to thermal-induced unfolding and, very likely, to cleavage of covalent bonds (Volkin *et al.*, 1991; Garza-Ramos *et al.*, 1994; Shoichet *et al.*, 1995).

2.1.4C. The Cost of Catalysis

Several enzymes with markedly different structure and function have been shown to exhibit a higher thermostability in low water environments compared to their stability in aqueous systems. However, under these conditions the enzymes express low catalytic rates. A higher capacity to carry out catalysis exists for an enzyme as the amount of water in contact with the enzyme is progressively increased because an enhancement of protein flexibility occurs. Thus, the cost of high catalytic rates is paid by a higher sensitivity of the enzyme to thermal denaturation.

2.1.5 Trapping of Enzyme Conformers in Low Water Systems

On the basis that water is an essential component of conformational mobility, it is theoretically possible to stabilize different protein conformations by adjustment of the amount of water in contact with the enzyme. Data obtained with enzymes dispersed in organic solvents or entrapped in reverse micelles show that two types of conformations may be stabilized in low water systems, namely one which corresponds to enzymes that exhibit particular catalytic properties (see below) and the other corresponds to intermediates of biological reactions that in water have relatively short life times (see Chapter 6, Section 3B.1.5).

Russell and Klibanov (1988) made the following observation. Subtilisin was lyophilized in the presence of a reversible active site inhibitor from aqueous solution, and the powder was washed with an organic solvent that effectively removed the inhibitor. The capacity of the enzyme to catalyze a transesterification reaction in octane was measured. Compared to the control (i.e., subtilisin not exposed to the inhibitor), subtilisin that had been exposed to the inhibitor exhibited substantially higher activity if the water content of the system was kept below 0.012%. The results were interpreted as

locking of an active conformation ("enzyme memory") of subtilisin that had been induced by the competitive inhibitor.

Stähl *et al.* (1990, 1991) followed the same rationale and incubated α-chymotrypsin with the anomalous substrate *N*-acetyl-D-Trp in aqueous media and precipitated it with propanol. The enzyme was dried and its capacity to form *N*-acetyl-D-Trp-OEt (from *N*-acetyl-D-Trp and ethanol) in cyclohexane with increasing water concentrations was determined. The enzyme catalyzed the synthesis of the D-ester in a bell shape curve with a maximum at 1.5 mM water; with 4 mM water very little synthesis of the D-ester was observed. The authors stated that the active site was "bio-imprinted" to catalyze the synthesis of the D-ester, a bio-imprinting that persisted as long as water was kept at low levels. In contrast, the enzyme imprinted with *N*-acetyl-L-Trp did not exhibit the anomalous activity; however it effectively catalyzed the synthesis of *N*-acetyl-L-Trp-OEt at rates that increased with water concentration. The increased flexibility induced by increasing water concentrations leads to an increase of "normal" catalysis, but it also provokes loss of the imprinting that allowed the enzyme to catalyze the "normal" reaction.

2.1.6 Structure of Enzymes in Organic Solvents

Research is ongoing to determine whether an enzyme acquires different structural features in different solvents that could account for differences in catalysis. Burke *et al.* (1989) examined the hydrogen-bonding of an active site histidine (labeled with ^{15}N) of α-lytic protease by high-resolution solid-state CP/MAS NMR in water and organic solvents; the tautomeric structure and hydrogen-bonded network at the active site was probed. Spectra revealed that the organic solvent did not perturb the geometrical

arrangements of the active site residues. This was the first direct physical evidence that enzymes do not denature when suspended in neat organic solvents. Burke *et al.* (1992) also examined the active center of α-chymotrypsin by solid-state ^{13}C-NMR spectroscopy after lyophilization and after suspending the powder in organic solvents. Lyophilization was observed to disrupt the active centers of a portion of the enzymes; further disruption occurred when the enzyme was suspended in water-stripping anhydrous hydrophilic solvents, but not when suspension was carried out with hydrophobic solvents. For example, dehydration leaves the hydrophobic binding pocket of α-chymotrypsin intact, but can cause enough distortion of the catalytic center to destroy its transition state-stabilizing power. Furthermore, unfolding can disrupt the active site of enzymes, which occurs in organic solvents when the enzyme acquires sufficient flexibility to escape its kinetically-trapped conformation in the lyophilized powder (Burke *et al.*, 1992).

Dong *et al.* (1996) determined the secondary structure of subtilisin and α-chymotrypsin by infrared spectroscopy; compared to the enzymes in solution, enzymes lyophilized and suspended in organic solvents showed an altered secondary structure. Griebenow and Klibanov (1995, 1997) also have shown by FT-IR spectroscopy that lyophilization dehydrates enzymes to the point of disrupting their active sites. Dehydration-induced secondary structural changes in lyophilized subtilisin were observed (decrease in α-helix content; increase in β-sheet content) compared to subtilisin dissolved in water. FT-IR data also revealed that active center blockage (some active sites are sterically shielded by neighbouring molecules) occurs upon lyophilization, which leads to accessibility limitations for the substrate, and consequently leads to a decreased enzymatic activity when suspended in organic solvents. The latter was

inferred from an increase in intermolecular β-sheet structures as a result of increased protein-protein contacts upon dehydration. Since enzymes in water have conformational dynamics and exist in several conformations, upon lyophilization the enzyme population consists of several kinetically-trapped conformations, which leads to a decrease in intact active centers, and hence, a decrease in enzymatic activity when suspended in organic solvents. The difference in intact active centers between lyophilized enzyme powder in organic solvents and the aqueous solution from which it came was shown to be only 2-fold by solid-state NMR (Burke et al., 1992). Hence, steric blockage and disruption of intact active centers due to dehydration upon lyophilization is responsible for only a small fraction of the loss of enzymatic activity in organic solvents compared with the situation in water.

A precise molecular explanation for the particular characteristics that enzymes acquire in organic solvents is lacking and further research to probe the molecular structure of enzymes in organic solvents is required. This is especially needed when dealing with changes in substrate specificity (Westcott and Klibanov, 1994; Zaks and Klibanov, 1986; Gololobov *et al.*, 1992; Calvet *et al.*, 1993; Tawaki and Klibanov, 1992; Stähl *et al.*, 1990, 1991), and enzyme memory (Russell and Klibanov, 1988; Zaks and Klibanov, 1988) or bioimprinting (Stähl *et al.*, 1991).

2.1.7 Crystal Structures

It is not feasible to crystallize enzymes from apolar solvents, as they are not soluble in such solvents. Enzymes crystallized from aqueous medium are therefore necessarily soaked in organic solvents. Fitzpatrick *et al.* (1994) showed that exposure of crystals of subtilisin Carlsberg to 50% acetonitrile led to cracking. Therefore, the authors

carried out crosslinking of subtilisin that had been crystallized in water, followed by soaking in anhydrous acetonitrile (Fitzpatrick *et al.*, 1993, 1994). The crystals were stable and diffracted at a resolution of 2.3 Å. This methodology was also used to ascertain the influence of acetonitrile on the crystal structure of crosslinked porcine pancreatic elastase (Allen *et al.*, 1996). In contrast to subtilisin and elastase, Yennawar *et al.* (1994, 1995) found that non-crosslinked crystals of γ-chymotrypsin were stable in hexane.

The data of Fitzpatrick *et al.* (1993, 1994) showed that the polypeptide chain of subtilisin in acetonitrile and water is markedly similar; some minor shifts of active site residues occurred. It was also observed that only a portion of the water molecules were removed by acetonitrile, and that acetonitrile occupied positions in which no water was visualized in crystals that had not been exposed to acetonitrile. Schmitke *et al.* (1997) examined the crystal structure of subtilisin in the presence of dioxane and compared it with that in acetonitrile. The crystal structures in acetonitrile, dioxane and water were markedly similar. In view of the fact that subtilisin exhibits distinct catalytic activities in different solvents (Ke *et al.*, 1996; Westcott *et al.*, 1996), it is significant that only minor changes were detected at the active site. Dioxane molecules were also observed to interact with the protein, and in comparison to the data with acetonitrile, different water molecules were replaced by two solvent molecules.

Allen *et al.* (1996) studied crystallized elastase and showed that soaking with acetonitrile did not modify the overall structure of the enzyme. They observed two out of nine acetonitrile molecules near the active site.

Yennawar *et al.* (1994, 1995) studied whether there are structure alterations in γ-chymotrypsin induced by hexane that could account for changes in substrate specificity (Zaks and Klibanov, 1986; Stähl *et al.*, 1990, 1991; Tawaki and Klibanov, 1992). The backbone of γ-chymotrypsin in hexane showed no major differences with the enzyme in water. However, some side-chains underwent rearrangements, particularly in the regions where hexane molecules appeared. Hexane molecules were found to replace some water molecule, but other hexane molecules appeared in the neighborhood of hydrophobic patches. Strengthening of H-bonds and electrostatic interactions was observed, which could induce rigidity of the enzyme molecule. The catalytic center showed no important differences with the enzyme in water.

The present crystallographic data indicates that organic solvents do not produce large modifications of the three-dimensional structure of the enzyme, nor of its catalytic center. Nevertheless, organic solvents produce changes in the orientation of some residues in the surface of the protein. The proteins appear to be more rigid in organic solvents than in water, which would be in agreement with data on the activity and thermostability of enzymes in organic solvents.

2.2 Nonaqueous Chemical Modification of Proteins: A Viable Approach to Protein Structure-Function Studies

Chemical modification was one of the first and most useful methods for studying structure-function relationships in proteins. Comprehensive descriptions of the techniques, reagents and strategies are available in general reviews of the field (Means and Feeney, 1971; Glazer *et al.*, 1976; Lundblad and Noyes, 1984; Imoto and Yamada, 1989; Lundblad, 1995). The properties of enzymes in organic solvents, as overviewed above, provided a rationale to study structure-function relationships via chemical

modification under nonaqueous conditions (Taralp, 1997, Doctoral Thesis; Taralp and Kaplan, 1997). Most importantly, lyophilized enzymes in organic solvents retain an active conformation that is similar to the native structure, which appears to be due to the maintenance of the essential water layer surrounding the protein molecule in the nonaqueous environment (Zaks and Klibanov, 1988; Klibanov, 1989; Chen and Sih, 1989; Westcott and Klibanov, 1994; Griebenow and Klibanov, 1995, 1997). Moreover, crystallized proteins have been observed to have essentially the same structure when placed in organic solvent as in water (Fitzpatrick *et al.*, 1993; Fitzpatrick *et al.*, 1994). If lyophilized proteins dispersed in organic solvents retain at least the essential elements of their native structure in aqueous solution, then the reactivity of their functional groups should reflect that in solution and provide information on the solution structure and its properties.

Taralp and Kaplan (1997) have demonstrated the feasibility and utility of nonaqueous chemical modification of proteins. Lyophilized proteins were reacted *in vacuo* with a volatile reagent or dispersed in octane and reacted with dissolved reagent. The greatly enhanced thermostability of enzymes due to their rigidity in organic solvents permitted the use of elevated temperatures to accelerate the chemical modification without unfolding or degrading the native structure (Zaks and Klibanov, 1988; Broos *et al.*, 1995). For the *in vacuo* approach, the derivatization is accomplished in a sealed reaction vessel under vacuum in which only the lyophilized protein and reagent are present. Since lyophilized proteins retain their structure and biological activity, the *in vacuo* reactivity of the various functional groups of the protein and the biological

activity of the *in vacuo* derivatized protein should reflect structure-function relationships in the protein.

There are significant potential advantages to carrying out chemical modification of proteins under *in vacuo* conditions (Taralp and Kaplan, 1997; Taralp, Doctoral Thesis, 1997); some are described as follows. (a) Water is not present as a competing reactant, which allows for virtually 100% reaction specificity with the protein and complete recovery of unreacted reagent. This is of considerable practical value in the utilization of very expensive reagents (e.g., radiolabeled reagents). (b) Derivatives that are unstable in water can be prepared (e.g., anhydrides of carboxyl groups). For example, bovine serum albumin was acetylated *in vacuo* with ethoxyformic anhydride and coupling to the activated carboxyl groups was carried out with a fluorescent amine in *N,N*-dimethylformamide (Taralp, 1997, Doctoral Thesis). The *in vacuo* procedure allowed for relatively small amounts of activating reagent to be used compared to the high concentrations required in the usual aqueous protocols for activating carboxyl groups. (c) A much greater diversity of derivatizing reagents can be employed. These include: reagents which react too rapidly with water to be used in the aqueous procedures for the derivatization of proteins; weak electrophiles such as alkyl halides or insoluble compounds which react too slowly to be practical derivatizing reagents in aqueous media. Their rate of reaction *in vacuo* can be accelerated at high temperatures since lyophilized proteins under vacuum are stable at temperatures well above 100°C. (d) No laborious purification of the derivatized protein is required since there is no solvent and any excess reagent(s) can be trapped out. (e) Much greater specificity towards the various classes of functional groups and individual functional groups within a

class can be achieved. Since lyophilized proteins have a pH memory (i.e., ionizable groups retain the degree of ionization they had in the solution from which they were prepared), maximum specificity or selective modification can be achieved by setting the pH at its optimum value for reaction of the ionizable groups of interest. The yield of a derivatized functional group is directly related to its pK_a, its surface exposure, and the pH of the solution from which the protein was lyophilized (LpH value). Thus, the physicochemical factors that govern the reactivity of protein functional groups in nonaqueous environments appear to reflect the protein solution structure prior to lyophilization (Taralp and Kaplan, 1997; Vakos *et al.*, 2000). The work of Taralp and Kaplan (1997; Taralp, Doctoral Thesis, 1997) has provided novel theoretical insights and has led to practical applications in the study of proteins (Taralp, Doctoral Thesis, 1997; Vakos *et al.*, 2000; Chapters 2-5).

Chemical modification of proteins was carried out using ^{13}C-iodomethane as a sensitive NMR probe for studying the chemical reactivity of functional groups in proteins under nonaqueous conditions (Taralp and Kaplan, 1997). Iodomethane has been marginally used for protein derivatization. Reductive dimethylation (using formaldehyde and borohydride) of amino groups in protein has been used in structure-function studies with retention of bioactivity in many cases (Means and Feeney, 1968, 1971, 1995). The modification is of benign nature and preserves protein structure. Proteins have complicated three-dimensional structures in which some amino acid residues are solvent exposed and others are buried. The functional groups with the highest probability of reaction are located at the surface of the protein where they are readily available to reagent (Means and Feeney, 1971). Met-29 of ribonuclease A was most easily

methylated by iodomethane, as it is solvent exposed and readily accessible to reagent, as evidenced by X-ray crystallography (Link and Stark, 1968; Richards and Wyckoff, 1971). Other proteins (basic myelin protein, basic pancreatic trypsin inhibitor, α-chymotrypsin) have been studied by the *S*-methyl approach (Jones *et al.*, 1975, 1976; Deber *et al.*, 1978; Harina *et al.*, 1978, 1980; Jaeck and Benz, 1979; Smith *et al.*, 1980).

In vivo, post-translational methylation of amino acid residues of various proteins appears to be essential in some cases to maintain biological function (Paik and Kim, 1980). Biological methylation reactions as typically carried out by highly specific *S*-adenosyl methionine methyl transferase enzymes (DiMaria *et al.*, 1982; Rowe *et al.*, 1986). Examples of biologically methylated proteins include histones, myosins, actins, cytochrome c, cytochrome c557, and ribosomal proteins to name a few (Paik and Kim, 1980). Sites of derivatization in the protein include lysyl, arginyl, histidyl, prolyl, glutaminyl, alanyl, methionyl, glutamyl, and aspartyl residues.

The fact that iodomethane has been used in few structure-function studies is possibly due to its limited solubility in water or its low reactivity compared to other reagents (e.g., acetic anhydride); advantage can be taken of the latter two properties since mild reagents are often preferred to carry out benign or conservative protein modifications. Since iodomethane is small and uncharged, its reaction with proteins should be unaffected by electrostatic interactions and minimally by steric factors.

Charged interactions of protein are partially responsible for the maintenance of protein structure. Native structure may be perturbed if there is an alteration in the original charge of a functional group. Moreover, the manner in which protein residues pack together ultimately determines protein structure. Chemical derivatization of protein

consequently leads to a change of steric bulk in the area immediately surrounding the modified functional group. A large group may not be considered bulky if there is nothing proximal to interact with it. Substitutions involving changes of steric bulk in proteins are often more serious within the protein core where residues are closely associated (Serrano *et al.*, 1992). Nevertheless, replacements where a small group is replaced by a larger group at the surface of a protein can still potentially have more serious consequences than a more isosteric replacement. The methyl group, which typically replaces a proton and some of its associated hydration shell, represents a small steric change, and while there are enzymes that can differentiate such minimal replacements (DiMaria *et al.*, 1982; Rowe *et al.*, 1986), the general effect of the modification on proteins is presumably small.

Advantage can be taken of the reactivity of iodomethane towards distinct functional groups to effect different modifications under appropriate conditions. In fact, the innovative studies of Taralp and Kaplan (1997) have shown that iodomethane forms stable derivatives with amino, phenolic hydroxyl, thiol, and imidazole nitrogen functional groups. The authors did not expect significant reaction of protein carboxyl groups with iodomethane due to their low nucleophilicity (although, iodomethane has been shown to methylate carboxyl groups in organic syntheses (March, 1985)), and therefore, did not take this into account in the interpretation of the results. Further investigations have shown that protein carboxyl groups actually do react with iodomethane *in vacuo* (Chapters 2 and 3; Vakos *et al.*, 2000).

Amino groups (α and ε) and the imidazole function of His residues can be derivatized (trimethylamino and dimethylimidazolium cation) to form a permanent positive charge on the side-chain at all pH values, which can be exploited in structure-

function studies. The pH memory effect was observed to be evident for the reactions of the ε- and α-amino groups and imidazole groups with iodomethane. At LpH 7.5, the degree of reaction of the ε-amino groups (pK_a 10-11), relative to that of the α-amino groups (pK_a 7-8), was much less than at LpH 10, while the imidazole group (pK_a 6-7) (Creighton, 1993) showed a much smaller difference. The results reflected the fact that in nonaqueous media there is no water present for a dynamic equilibrium to be established between the two ionization states, i.e. a group that is protonated will not deprotonate and its rate of reaction will be negligible relative to the deprotonated group. Thus, since functional groups in proteins have pK_a values varying from 3.5 to 12, the pH memory phenomenon has the potential to be utilized as a means of achieving selective chemical modifications of ionizable functional groups by controlling the pH of lyophilization (LpH).

An enhanced ability over the aqueous protocols to probe protein structure was observed with the *in vacuo* ^{13}C-methylation of proteins (Taralp, 1997, Doctoral Thesis). A notable difference in the reactivity of the tyrosine phenolic hydroxyl function was observed between insulin, α-chymotrypsin, and ribonuclease A in a nonaqueous environment in that only in insulin was a ^{13}C-NMR resonance for the methylated derivative of this functional group observed at LpH 10. Under aqueous conditions the tyrosine phenolic hydroxyl function is readily methylated at pH 10 with these model proteins. These results suggest that the phenolic side-chains are buried in the major conformational states of α-chymotrypsin and ribonuclease A in solution at pH 10. These side-chains react in an aqueous but not in a nonaqueous environment because of the dynamic equilibrium that exists between the various conformational states in solution.

Although the tyrosyl side-chains are exposed only in minor conformations, this could lead to substantial reaction over a period of time due to Le Châtelier's principle. No such dynamic equilibria exist in the nonaqueous state and no substantial modification can occur. Insulin, being a very small protein, is unfolded to a large extent in solution at pH 10, and when a sample is prepared at LpH 10, some or all of the tyrosine side-chains are exposed.

^{13}C-Methylation of α-amino groups results in the trimethylamino derivative of those groups, which are evidenced on the ^{13}C-NMR spectrum as distinct resonances. If a protein has one, two or three polypeptide chains, it will have the respective number of resonances corresponding to the number of α-amino-termini present in addition to the resonance corresponding to the trimethyl ε-amino groups of Lys residues, which overlap and appear as one resonance. An additional peak to the number of expected trimethyl α-amino acid residues is indicative of proteolytic breakdown of the protein of interest or of the presence of minor impurities. The latter are usually evidenced as small resonances compared to the major peaks. Chemical modification of proteins in aqueous media often disrupts the native structure and may lead to degradation by trace amounts of proteolytic enzymes. ^{13}C-Methylated α-chymotrypsin revealed additional resonances corresponding to trimethyl α-amino acid residues. Most model proteins that were methylated *in vacuo* gave the expected number of resonances for the trimethyl α-amino acid residues. Thus, the ability to detect free α-amino-termini demonstrates that the nonaqueous ^{13}C-methylation procedure with ^{13}C-iodomethane could provide a means to determine the homogeneity of protein preparations. Moreover, the absence of a resonance for a

trimethyl α-amino residue indicates that the N-terminus is blocked (Taralp, 1997, Doctoral Thesis; Chapter 2, Vakos *et al.*, 2000).

Taralp (1997, Doctoral Thesis; Stewart *et al.*, 1997) showed that *in vacuo* ^{13}C-methylated α-chymotrypsin is inactivated via the methylation of the active site His-57 residue. Ligands that bind to proteins have been shown to protect functional groups in the binding regions from chemical modification (Means and Feeney, 1971; Bosshard, 1979). However, this approach for identifying functional groups in the active sites of enzymes has found only limited application. Since dynamic binding equilibria exist in aqueous media, and most competitive inhibitors have relatively weak binding affinities, active site functional groups are still readily modified even in the presence of these inhibitors. In contrast, under the *in vacuo* nonaqueous conditions, dynamic equilibria cannot exist and if nonvolatile ligand is bound it cannot dissociate and the residues in the binding site would not be expected to react. α-Chymotrypsin was lyophilized from a solution containing 10 mM competitive inhibitor (e.g., indole, K_I=0.72 mM; *N*-acetyl-L-Trp, K_I=17.5 mM; Barman, 1969). Indole, a good competitive inhibitor, did not offer any protection, whereas *N*-acetyl-L-Trp, a weaker inhibitor, gave substantial protection. With the ligand present at 10 mM, it is expected that virtually all the indole binding sites in the lyophilized enzyme are occupied, whereas a substantial portion of the binding sites are free in the case of the *N*-acetyl-L-Trp. The *N*-acetyl-proprionyl moiety of the latter covers, and hence protects, the catalytic site. Importantly, such information would be very difficult to obtain in an aqueous medium and provides yet another example of the potential utility of the chemical modification of proteins in

nonaqueous environments. Stewart *et al.* (1997) have carried out detailed protection studies with α-chymotrypsin.

3. Overview of the present study

The work presented in this thesis illustrates the applicability of low water systems [namely the *in vacuo* chemical modification technique (Chapters 2-5) and the reverse micellar systems (Chapter 6)] to the study of proteins in order to gain insight into protein structure and function. An outline of each chapter is presented below and some of the novel findings are described. The applications that have stemmed from this body of research are currently being pursued in our laboratory and already have been proven viable in the characterization of proteins (i.e., for N- and C-terminal identification).

Chapter 2: Use of the pH Memory Effect in Lyophilized Proteins to Achieve Preferential Methylation of α-Amino Groups

Taralp and Kaplan (1997) have demonstrated the utility of iodomethane as a protein-modifying reagent. These authors directly observed the reaction of functional groups in protein with ^{13}C-iodomethane under nonaqueous conditions by ^{13}C-NMR spectroscopy. A feature of this nonaqueous modification procedure was that α- and ε-amino groups were converted to the ^{13}C-trimethylammonium derivatives, which greatly enhanced the sensitivity of detection of individual groups by ^{13}C-NMR spectroscopy. In the present investigation, the nonaqueous methylation procedure was combined with the current NMR technology to yield a sensitive method for analyzing the presence of α-amino groups in protein preparations. The procedure allows for the identification of the number and types of N-terminal ^{13}C-trimethylamino derivatives of amino acid residues derived from ^{13}C-methylated intact protein or from free amino acids derived

from methylated protein hydrolysates based on the ^{13}C-chemical shifts of ^{13}C-trimethyl amino acid or di- or tripeptide standards. The method is facile and can be employed as an analytical tool to investigate the homogeneity of a protein preparation and to verify whether a protein has a free or blocked amino-terminus. The pH memory effect (Zaks and Klibanov, 1988) can be used to control the ionization state of amino groups in lyophilized proteins, and hence their chemical reactivity towards modifying reagents. Advantage was taken of this property to achieve preferential or selective modification among functional groups in protein, in particular groups of the same class. In the present work, preferential trimethylation of α-amino groups without substantial derivatization of ε-amino groups was demonstrated. Protein lyophilized from aqueous solution at pH 6 or 7 (LpH 6-7) was reacted *in vacuo* with ^{13}C-iodomethane at 75°C. Minor resonances of free α-amino groups arising from proteolysis of the major protein or from protein impurities were detectable. SDS-PAGE cannot distinguish subtle differences in contaminating proteins of similar size or heterogeneity. *In vacuo* reaction with ^{13}C-iodomethane also resulted in the formation of ^{13}C-methyl esters of the protein carboxyl groups, the resonances of which were removed by hydroxylamine or base and heat to clarify the spectra, leaving only the resonances arising from the ^{13}C-trimethylamino groups. Both ^{13}C and ^{14}N NMR spectroscopy (Chapters 2 and 4) were utilized to characterize the trimethylammonium derivatives of α- and ε-amino groups formed by reaction with iodomethane (^{13}C, ^{12}C).

Chapter 3: *In Vacuo* Esterification of Carboxyl Groups in Lyophilized Proteins

The utility of ^{13}C-iodomethane as a carboxyl group-modifying reagent was investigated. As mentioned above, carboxyl ^{13}C-methyl ester derivatives of protein

carboxyl functions were not observed with the *in vacuo* methylation technique in the studies of Taralp and Kaplan (1997). It is demonstrated that the pH memory effect of lyophilized proteins (Zaks and Klibanov, 1988) can be utilized to effect preferential esterification of protein carboxyl groups at low LpH values (between 3-4). An increase in the extent of methylation was observed with an increase of deprotonated carboxyl groups at higher LpH values, but trimethylation of amino groups was the predominant reaction at LpH values greater than 4 or 5. The current NMR technology provides sufficient sensitivity for the identification of the ^{13}C-chemical shifts of the different types of carboxyl methyl esters that can be formed by reacting protein with methylating reagents. An in-depth study of several proteins and appropriate standard compounds permitted the latter. Reaction with iodomethane was useful in that the pH memory effect of lyophilized proteins (Zaks and Klibanov, 1988; Taralp, Doctoral Thesis, 1997) was confirmed; but the results show that when alkylating amino groups (Chapter 2), *in vacuo*-synthesized esters of carboxyl groups will be formed and should be removed.

In contrast to the *in vacuo* reaction with ^{13}C-iodomethane, the novel *in vacuo* esterification reaction with gaseous ^{13}C-methanol/HCl or ^{13}C-ethanol/HCl at 75 °C resulted in the exclusive formation of α-, γ- and δ-carboxyl ^{13}C-methyl or ^{13}C-ethyl ester derivatives, which were distinguishable by their distinct chemical shifts. To our knowledge, the *in vacuo* esterification chemical modification, as carried out on lyophilized protein under nonaqueous conditions, has not been performed previously. Resonances from individual esterified side-chains of aspartic (γ) or glutamic (δ) acid residues could not be greatly resolved. Significantly, the chemical shift(s) of the C-terminal α-carboxyl group(s) showed a strong dependence on the nature of the

C-terminal amino acid and sequence (i.e., penultimate amino acid residue). It was possible to determine the number of C-termini and make a tentative identification of the C-terminal amino acid for several proteins by comparison of the chemical shifts of the resonances of appropriate standard compounds. The extent of derivatization of protein carboxyl groups with iodomethane or acidic alcohol can be controlled (extent depends on reaction time, LpH, pK_a and extent of surface exposure of functional groups) and exploited in structure-function or protein characterization studies, as the protein can retain its structural integrity in the lyophilized state throughout the *in vacuo* procedure. This study has provided further supporting evidence for the general applicability of the *in vacuo* nonaqueous methodology to effect the chemical modification of proteins.

Chapter 4: The Use of ^{14}N NMR Spectroscopy To Study Trimethylated Amino Groups in Methylated Protein Hydrolysates

Successfully adapting NMR techniques for use in the area of protein characterization is an ongoing theme in protein chemistry. The use of ^{14}N NMR spectroscopy for elucidating protein structure-function relationships has been limited (Zhang *et al.*, 1994). For example, amino group ^{14}N resonances of protein are usually very broad because ^{14}N nuclei are quadrupolar and have very efficient relaxation (quadrupolar relaxation mechanism) (Witanowski and Webb, 1972). Chemical modification of α- and ε-amino groups of amino acids in proteins with iodomethane leads to quaternized amino functional groups in the form of trimethylammonium derivatives, which possess exceptional electronic symmetry (dipolar relaxation mechanism with longer relaxation times and sharper resonances). The methylated amino groups of protein or protein hydrolysates are readily observed as very sharp resonances by ^{14}N NMR spectroscopy. Detection of the methylated residues by this simple approach is further

simplified by carrying out the NMR analyses at high pH, where other nitrogen resonances are too broad to be observed. The ability to characterize α- and ε-amino groups in this manner has potential applications in structure-function studies, as well as practical uses in protein characterization such as in assessing protein homogeneity by identifying α-amino groups. Moreover, iodomethane provides a unique opportunity to directly observe and investigate the reaction, local environment and motion of amino groups in any protein by ^{14}N NMR spectroscopy.

Chapter 5: High Specific Activity of Tracer Molecules Prepared *In Vacuo* with ^{14}C-Iodomethane

Implementation of the nonaqueous chemical modification methodology of Taralp and Kaplan (1997) affords the possibility to radiolabel proteins with high specific radioactivity using ^{14}C- and ^{3}H-enriched iodomethane. Such high levels of specific activity cannot be achieved if the chemical modification is carried out in aqueous environments with these nuclei due to the competition of water with the protein for the reagent, which can be very expensive. The chemical modification of lyophilized proteins under nonaqueous conditions can maintain the essential elements of the native protein structure. Moreover, the pitfalls of aqueous modification, namely proteolysis and denaturation can be avoided, which is important if radiolabeled protein is used in structure-function studies. A typical aqueous strategy is to use radiolabeled and unradiolabeled reagent in excess, which can be very expensive and lead to extensive modification and not a very radioactive protein derivative with a decrease in bioactivity. Thus, in the nonaqueous chemical modification approach proteins that are not enzymes can be chemically tagged in a manner that can serve as a probe of structure and function. [Simply carrying out an activity assay can test the maintainance of activity of a

chemically modified enzyme, but the function of a protein that is not an enzyme is more difficult to assess.] Radiolabeled model proteins (e.g., *Bacillus thuringiensis* insecticidal toxin, insulin) can be used in tracer studies to elucidate their mechanisms of action.

Chapter 6: Cholesterol Esterase: A. Purification, B. Characterization of Taurocholate-Induced Dimerization in Sodium Dodecylsulfate, C. Bile Salt-Modulated Stereoselection in the Hydrolysis of α-Tocopheryl Acetates; Effect of Ester Group Chain Length on the Rates of Hydrolysis of Synthetic α-Tocopheryl Esters

The first objective of the present study of CE was to purify bovine and porcine cholesterol esterase (BCE, PCE) and human milk bile salt-activated lipase (BAL) using a cholate-derivatized affinity column and a taurocholate concentration gradient in the eluent (Moore *et al.*, 1996). An unusual phenomenon was observed on SDS-PAGE in the course of purifying BCE. Taurocholate induces dimerization of BCE in SDS. The fact that such a specific interaction does occur under denaturing conditions suggests that under non-denaturing conditions bile salts induce oligomerization of BCE.

The second objective was to elucidate the factors that influence the apparent dimerization of CE on heating in SDS and to characterize the nature of the dimerization. Protein-protein interactions can be studied in reverse micelles and information observable by this approach is highlighted for prospective studies.

The third objective was to study the bile salt-modulated stereoselectivity of CE-catalyzed hydrolysis of α-tocopheryl acetates. In competitive experiments, the diastereoselectivity of CE depends on the bile salt used, which is relevant if the bile salt-modulated effect on CE-catalyzed reactions is to be exploited in organic syntheses. A preliminary mechanistic study of CE was carried out with synthetic α-tocopheryl ester substrates (*RRR*-α-T esters) of varying chain length. The effect of the chain length on

reactivity reflects the lifetime of the acylenzyme intermediates as well as product inhibition.

4. References

Adlercreutz, P. (1991) *Eur. J. Biochem.*, **199:** 609-614.

Affleck, R., Haynes, C. A., and Clark, D. S. (1992a) *Proc. Natl. Acad. Sci. U.S.A.*, **89:** 5167-5170.

Affleck, R., Xu, Z.-F., Suzawa, V., Focht, K., Clark, D. S., and Dordick, J. S. (1992b) *Proc. Natl. Acad. Sci. U.S.A.*, **89:** 1100-1104.

Allen, K. N., Bellamacina, C. R., Ding, X., Jeffery, C. J., Mattos, C., Petsko, G. A., and Ringe, D. (1996) *J. Phys. Chem.*, **100:** 2605-2611.

Almarsson, O. and Klibanov, A.M. (1996) *Biotechnol. Bioeng.*, **49:** 87-92.

Asther, M. and Meunier, J. C. (1990) *Enzyme Microb. Technol.*, **12:** 902-905.

Ayala, G., Tuena de Gómez-Puyou, M., Gómez-Puyou, A., and Darszon, A. (1986) *FEBS Lett.*, **203:** 41-43.

Bell, G., Halling, P.J., Moore, B.D., Partridge, J., and Rees, D.G. (1995) *Trends Biotechnol.*, **13:** 468-473.

Biemann, K. (1992) In *Annu. Rev. Biochem.*, **61:** 977-1010.

Bone, S. (1987) *Biochim. Biophys. Acta*, **916:** 128-134.

Bone, S. and Plethig, R. (1982) *J. Mol. Biol.*, **157:** 571-575.

Bosshard, H. R. (1979) *Methods Biochem. Anal.*, **25:** 273-301.

Boyer (1970) *The Enzymes*, 3rd edn., Academic Press, **1**.

Boyer (1971) *The Enzymes*, 3rd edn., Academic Press, **3**.

Broos, J., Sakodinskaya, I.K., Engbersen, J.F.J., Verboon, W., and Reinhoudt, D.N. (1995) *J. Chem. Soc., Chem. Commun.*, pp. 255-256.

Bru, R., Sánchez-Ferrer, A., and Garcia-Carmona, F. (1995) *Biochem. J.*, **310:** 721-739.

Brzozowski, A. M., Derewende, U., Derewende, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Huge-Jensen, B., Patkar, S. A., and Thim, L. (1991) *Nature*, **351:** 491-494.

Burke, P.A., Smith, S.O., Bachovchin, W.W., and Klibanov, A.M. (1989) *J. Am. Chem. Soc.*, **111:** 8290-8291.

Burke, P.A., Griffin, R.G., and Klibanov, A.M. (1992) *J. Biol. Chem.*, **267:** 20057-20064.

Burke, P.A., Griffin, R.G., and Klibanov, A.M. (1993) *Biotechnol. Bioeng.*, **42:** 87-94.

Calvet, S., Clapes, P., Torres, J. J., Valencia, G., Feixas, J., and Adlercreutz, P. (1993) *Biochim. Biophys. Acta*, **1164:** 189-196.

Campbell, I. D. and Dwek, R. A. (1984) *Biological Spectroscopy*, Benjamin/Cummings, Menlo Park, CA.

Chang, G.-G. and Shiao, S.-L. (1994) *Eur. J. Biochem.*, **220:** 861-870.

Chen, C.-S. and Sih, C. J. (1989) *Agnew. Chem. Int. Ed. Engl.*, **28:** 695-707.

Choma, C. T., Surewicz, W. K., Carey, P. R., Pozsgay, M. and Kaplan, H. (1990) *J. Protein Chem.*, **9**: 87-94.

Cleland, W. W. (1982) *Crit. Rev. Biochem.*, **13**: 385-428.

Clore, G. M. and Gronenborn, A. M. (1991a) *Annu. Rev. Biophys. Biophys. Chem.*, **20**: 29-63.

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

Cohen (1971) In *The Enzymes* (Boyer, ed.), 3rd edn., Academic Press, **3**.

Creighton, T. E. (1993). *Proteins: Structure and Molecular Properties*, 2nd ed., W.H. Freeman and Company.

Dabulis, K. and Klibanov, A. M. (1993) *Biotech. Bioeng.*, **41**: 566-571.

Darbre, A. (1986). *Practical Protein Chemistry* (Darbre, A., ed.), Wiley, New York.

Dastoli, F. R., Musto, N. A., and Price, S. (1966) *Arch. Biochem. Biophys.*, **115**: 44–47.

Dastoli, F. R. and Price, S. (1967) *Arch. Biochem. Biophys.*, **118**: 163–165.

Dastoli, F. R. and Price, S. (1967) *Arch. Biochem. Biophys.*, **122**: 289–291.

Deber, C. M., Moscarello, M. A. and Wood, D. D. (1978) *Biochemistry*, **17**: 898.

DiMaria, P., Kim, S., and Paik, W. K. (1982) *Biochemistry*, **21**: 1036-1044.

Dong, A., Meyer, J. D., Kendrick, B. S., Manning, M. C., and Carpenter, J. F. (1996) *Arch. Biochem. Biophys.*, **334**: 406-414.

Dordick, J. S. (1989) *Enz. Microb. Technol.*, **11**: 194-211.

Dordick, J. S. (1992) *Biotechnol. Prog.*, **8**: 259-267.

Dorovska-Taran, V., Veeger, C., and Visser, A. J. W. G. (1993) *Eur. J. Biochem.*, **218**: 1013-1019.

Drago, R. S. (1992) Physical Methods for Chemists, 2nd edn., Saunders College Publishing, New York.

Duggleby, R. G. and Kaplan, H. (1975a) *Biochemistry*, **14**: 5168-5175.

Duggleby, R. G. and Kaplan, H. (1975b) *Anal. Biochem.*, **65**: 346-354.

Escamilla, E., Contreras, M., Escobar, L., and Ayala, G. (1992) *Biological Electron Transport in Reverse Micellar Systems: From Enzymes to Cells in Biomolecules in Organic Solvents*, (Gómez-Puyou, A., Darszon, A., and Tuena de Gómez-Puyou, M., eds.), CRC Press, Boca Raton, FL., 219-241.

Evans, J. N. S. (1995) *Biomolecular NMR Spectroscopy*, Oxford University Press, New York.

Fernández-Velasco, D. A., Garza-Ramos, G., Ramirez, L., Shoshani, L., Darszon, A., Tuena de Gómez-Puyou, M., and Gómez-Puyou, A. (1992) *Eur. J. Biochem.*, **205**: 501-508.

Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd ed., W.H. Freeman and Company, New York.

Fitzpatrick, P.A., Steinmetz, A.C.U., Ringe, D., and Klibanov, A.M. (1993) *Proc. Natl. Acad. Sci. U.S.A.*, **90**: 8653-8657.

Fitzpatrick, P.A., Ringe, D., and Klibanov, A.M. (1994) *Biochem. Biophys. Res. Commun.*, **198**: 675-681.

Garza-Ramos, G., Darszon, A., Tuena de Gómez-Puyou, M. and Gómez-Puyou, A. (1989) *Biochemistry*, **28**: 3177-3182.

Garza-Ramos, G., Darszon, A., Tuena de Gómez-Puyou, M. and Gómez-Puyou, A. (1990) *Biochemistry*, **29**: 751-757.

Garza-Ramos, G., Darszon, A., Tuena de Gómez-Puyou and Gómez-Puyou, A. (1992) *Biochemistry*, **29:** 751-757.

Garza-Ramos, G., Tuena de Gómez-Puyou, Gómez-Puyou, A., and Gracy, R. W. (1992) *Eur. J. Biochem.*, **208:** 389-395.

Garza-Ramos, G., Tuena de Gómez-Puyou, M., Gómez-Puyou, M., Yuksel, U. and Gracy, R. W. (1994) *Biochemistry*, **33:** 6960–6965.

Glasel, J. A. and Deutscher, M. (1995) *Introduction to Biophysical Methods for Protein and Nucleic Acid Research*, Academic Press.

Glazer, A. N., Delange, R. J. and Sigman, D. S. (1976) In *Laboratory Techniques in Biochemistry and Molecular Biology* (Work, T. S. and Work, E., eds.) North-Holland Publishing Company, **4**.

Griebenow, K. and Klibanov, A.M. (1995) *Proc. Natl. Acad. Sci. U.S.A.*, **92:** 10969-10976.

Griebenow, K. and Klibanov, A.M. (1997) *Biotechnol. Bioeng.*, **53:** 351-362.

Gololobov, M. Y., Voyushina, T. L., Stepanov, V. M., and Adlercreutz, P. (1992) *FEBS Lett.*, **307:** 309-312.

Gómez-Puyou, A., Darszon, A., and Tuena de Gómez-Puyou, M. (1992) *Biomolecules in Organic Solvents*, CRC Press, Boca Raton, FL.

Gorman, L. A. S. and Dordick, J. S. (1992) *Biotechnol. Bioeng.*, **39:** 392-397.

Guinn, R. M., Skerker, P.S., Kavanaugh, P., and Clark, D.S. (1991) *Biotechnol. Bioeng.*, **37:** 303-308.

Gupta, M. N. (1992) *J. Biochem.*, **203:** 25-32.

Halling, P. (1994) *Enzyme Microb. Technol.*, **16:** 178-206.

Hames, B. D. and Rickwood, D. (1990) *Gel Electrophoresis of Proteins: A Practical Approach*, 2nd edn., IRL Press.

Harina, B. M., Dyckes, D. F., Willcott, M. R., and Jones Jr., W. C. (1978) *J. Am. Chem. Soc.*, **100:** 4897.

Harina, B. M., Dyckes, D. F., Willcott, M. R., and Jones Jr., W. C. (1980) *J. Am. Chem. Soc.*, **102:** 1120.

Harris, R. K. (1983) *Nuclear Magnetic Resonance Spectroscopy: A Physicochemical View*, Pitman Publishing Inc.

Hasnain, S., Visentin, L. P., and Kaplan, H. (1977) *Eur. J. Biochem.*, **80:** 35-41.

Herzberg, G. (1945) *Molecular Spectra and Molecular Structure. II. Infrared and Raman Spectra of Polyatomic Molecules*, Van Nostrand, Princeton, NJ.

Herzberg, G. (1950) *Molecular Spectra and Molecular Structure. I. Spectra of Diatomic Molecules*, Van Nostrand, Princeton, NJ.

Herzberg, G. (1966) *Molecular Spectra and Molecular Structure. III. Electronic Spectra and Electronic Structure of Polyatomic Molecules*, Van Nostrand, Princeton, NJ.

Hirs, C. H. W. (1967) *Meth. Enzymol.*, Academic Press, **11**.

Hirs, C. H. W. and Timasheff, S. N. (1972) *Meth. Enzymol.*, Academic Press, **25B**.

Hirs, C. H. W. and Timasheff, S. N. (1985) *Meth. Enzymol.*, Academic Press, **117**.
Imoto, T. and Yamada, H. (1989) In *Protein Function* (Creighton, T. E., Ed.), IRL Press, Oxford.

Hirs, C. H. W. and Timasheff, S. N. (1986a) *Meth. Enzymol.*, Academic Press, **130**.

Hirs, C. H. W. and Timasheff, S. N. (1986b) *Meth. Enzymol.*, Academic Press, **131**.

Holbrook, J. J., Liljas, A., Steindel, J., and Rossman, M. G. (1975) *Enzymes*, **11(part A):** 191-292.

Jaeck, G. and Benz, F. W. (1979) *Biochem. Biophys. Res. Commun.*, **86:** 885.

James, T. L. and Oppenheimer, N. J. (1994) *Meth. Enzymol.*, Academic Press, **239**.

Jansen, E. F., Nutting, M.-D. F., Jang, R. and Balls, A. K. (1949) *J. Biol. Chem.*, **179:** 189.

Jones, Jr., W. C., Rothgeb, T. M. and Gurd, F. R. N. (1975) *J. Am. Chem. Soc.*, **97:** 3876.

Jones Jr., W. C., Rothgeb, T. M. and Gurd, F. R. N. (1976) *J. Biol. Chem.*, **251:** 7452.

Karger, B. L. and Hancock, W. S. (1996a) *Meth. Enzymol.*, Academic Press, **270**.

Karger, B. L. and Hancock, W. S. (1996b) *Meth. Enzymol.*, Academic Press, **271**.

Karplus, M. and McCammon, J. A. (1983) *Ann. Rev. Biochem.*, **52:** 263-300.

Ke, T., Westcott, C. R., and Klibanov, A. M. (1996) *J. Am. Chem. Soc.*, **118:** 3366-3374.

Kernen, P., Agositi, R. D., Strasser, R. J., and Darszon, A. (1997) *Biochim. Biophys. Acta*

Kitaguchi, H. and Klibanov, A.M. (1989) *J. Am. Chem. Soc.* **111:** 9272-9273.

Khmelnitsky, Y. L., Hilhorst, R., Visser, A. J. W. G. and Veeger, C. (1993) *Eur. J. Biochem.*, **211:** 73-77.

Khmelnitsky, Y.L., Welch, S.H., Clark, D.S., and Dordick, J.S. (1994) *J. Am. Chem. Soc.*, **116:** 2647-2648.

Klibanov, A.M. (1989) *TIBS*, **14:** 141-144.

Klibanov, A.M. (1997) *Trends Biotechnol.*, **15:** 96-101; *and references therein.*

Koskinen, A.M.P. and Klibanov, A.M., eds. (1996) *Enzymatic Reactions in Organic Media*, Blackie Academic and Professional.

Leggett Bailey, J. (1962) *Techniques in Protein Chemistry*, Elsevier Publishing Company.

Link, T. P. and Stark, G. R. (1968) *J. Biol. Chem.*, **243:** 1082-1085.

Luisi, P. L., Giomini, M., Pileni, M. P., and Robinson, B. H. (1988) *Biochim. Biophys. Acta*, **947:** 209-246.

Lundblad, R. L. and Noyes, C. M. (1984) *Chemical Reagents for Protein Modification*, CRC Press, 1 and 2.

Lundblad, R. L. (1995) *Techniques of Protein Modification*, CRC Press, Boca Raton.

March, J. (1985) In *Advanced Organic Chemistry*, Third Edition, Wiley-Interscience, New York.

March, J. (1988) *Advanced Organic Chemistry: Reactions, Mechanisms and Structure*, 3rd edn., John Wiley and Sons.

Martinek, K., Klyachko, N. L., Kabanov, A. V., Khmelnitsky, Y. L., and Levashov, A. V. (1989) *Biochim. Biophys. Acta*, **981:** 161-172.

Marzola, P., Forte, C., Pinzino, C., and Veracini, C. A. (1991) *FEBS Lett.*, **289:** 29-32.

McCloskey, J. A. (1990) *Meth. Enzymol.*, Academic Press, **193**.

Means, G.E. and Feeney, R.E. (1968) *Biochemistry*, **7:** 2192-2201.

Means, G.E. and Feeney, R.E. (1971) *Chemical Modification of Proteins*, Holden-Day.

Means, G.E. and Feeney, R.E. (1990) *Bioconjug. Chem.*, **1:** 2-12.

Means, G.E. and Feeney, R.E. (1995) *Anal. Biochem.*, **224:** 1-16.

Moore, A.N.J., Vakos, H.T, Neville, T.A.M., Tonge, P., Arya, P., and Burton, G.W. (1996) *Biotechnology Techniques.* **10:** 523-528.

Mozhaev, V.V., Poltevsky, K.J., Slepnev, V.I., Badun, G.A., and Levashov, A.V. (1991) *FEBS Lett.*, **292:** 159-161.

Nicot, C. and Waks, M. (1996) *Biotechnol. Gen. Engin. Rev.*, **13:** 267-314.

Okahata, Y. and Ijiro, K. (1992) *Bull. Chem. Soc. Jpn.*, **65:** 2411-2420.

Paik, W. K. and Kim, S. (1980) *Protein Methylation*, Wiley-Interscience, New York.

Richards, F. M. and Wyckoff, H. W. (1971) In *The Enzymes*, 3rd edn., (Boyer, P. D., ed.), Academic Press, **3:** 683.

Paradkar, V.M. and Dordick, J.S. (1994) *J. Am. Chem. Soc.*, **116:** 5009-5010.

Pileni, M. P. (1989) *Structure and Reactivity in Reverse Micelles Studies in Physical and Theoretical Chemistry*, Elsevier, Amsterdam, **65.**

Profy, A. T. and Schimmel, P. (1988) *Prog. Nucleic Acid Res.*, **35:** 1-26.

Purich, D. L. (1980) *Meth. Enzymol.*, Academic Press, **64.**

Reid, C. and Rand, R. P. (1997) *Biophys. J.*, **71:** 2245-2255.

Reslow, M., Adlecreutz, P., and Mattiason, B. (1987) *Process Appl. Microbiol. Biotechnol.*, **26:** 1-8.

Roberts, G. C. K., ed. (1993) *NMR of Macromolecules: A Practical Approach*, Oxford University Press, New York.

Rowe, P. M., Wright, L. S. and Siegel, F. L. (1986) *J. Biol. Chem.*, **261:** 7060-7069.

Rupley, J. A., Gratton, E., and Careri, G. (1983) *TIBS*, **8:** 18-22.

Russell, A.J. and Klibanov, A.M. (1987) *J. Biol. Chem.*, **263:** 11624-11626.

Russell, A. J. and Klibanov, A. M. (1988) *J. Biol. Chem.*, **263:** 11624-11626.

Ryu, K. and Dordick, J. S. (1992) *Biochemistry*, **31:** 2588-2598.

Ryu, K. and Dordick, J. S. (1989) *J. Am. Chem. Soc.*, **111:** 8026-8027.

Sauer, K. (1995) *Meth. Enzymol.*, Academic Press, **246.**

Schmitke, J. L., Westcott, C. R., and Klibanov, A. M. (1996) *J. Am. Chem. Soc.*, **118:** 3360-3365.

Schmitke, J. L., Stern, L. J., and Klibanov, A. M. (1997) *Proc. Natl. Acad. Sci. U.S.A.*, **94:** 4250-4255.

Serrano, L., Kellis, J. T., Cann, P., Jr., Matouschek, A. and Fersht, A. R. (1992) *J. Mol. Biol.*, **224:** 783-804

Shoichet, B. K., Basse, W. A., Kuroki, R., and Matthews, B. W. (1995) *Proc. Natl. Acad. Sci. U. S. A.*, **92:** 452-456.

Smith, K. M., Kehres, L. A., and Tabba, H. D. (1980) *J. Am. Chem. Soc.*, **102:** 7151-7152.

Stähl, M., Mänsson, M.-O., and Mosbach, K. (1990) *Biotechnol. Lett.*, **12:** 161-166.

Stähl, Jeppsson-Wistrand, U., Mänsson, M.-O., and Mosbach, K. (1991) *J. Am. Chem. Soc.*, **113:** 9366-9368.

Strambini, G. B. and Gonnelli, M. (1988) *J. Phys. Chem.*, **92:** 2850-2853.

Stewart, N.A.S., Taralp, A., and Kaplan, H. (1997) *Biochem. Biophys. Res. Commun.*, **240:** 27-31.

Taralp, A. (1997) In *Aqueous and Nonaqueous Chemical Approaches to Elucidating Structure and Function in Proteins.* Ph.D. Thesis, Department of Chemistry, University of Ottawa, Ottawa, ON.

Taralp, A. and Kaplan, H. (1997) *J. Protein Chem.*, **16:** 183-193.

Tawaki, S. and Klibanov, A.M. (1992) *J. Am. Chem. Soc.*, **114:** 1882-1884.

Tuena de Gómez-Puyou, M. and Gómez-Puyou, A. (1998) *Crit. Rev. Biochem. Mol. Biol.*, **33:** 53-89.

Vakos, H. T., Kaplan, H., Black, B., Dawson, B., and Hefford, M. A. (2000) *J. Protein Chem.*, **19:** 231-7.

Volkin, D. B., Staubli, A., Langer, R., and Klibanov, A. M. (1991) *Biotechnol. Bioeng.*, **37:** 843-853.

Wangikar, P.P., Graycar, T. P., Estell, D.A., Clark, D.S., and Dordick, J.S. (1993) *J. Am.Chem. Soc.*, **115:** 12231-12237.

Wangikar, P. P., Rich, J. O., Clark, D. S., and Dordick, J. S. (1995) *Biochemistry*, **34:** 12302-12310.

Wangikar, P.P., Carmichael, D., Clark, D.S., and Dordick, J.S. (1996) *Biotechnol. Bioeng.*, **50:** 329-335.

Westcott, C.R. and Klibanov, A.M. (1993) *J. Am. Chem. Soc.*, **115:** 1629-1631.

Westcott, C. R. and Klibanov, A. M. (1994) *Biochim. Biophys. Acta*, **1206:** 1-9.

Westcott C. R., Noritomi, H., and Klibanov, A. M. (1996) *J. Am. Chem. Soc.*, **118:** 10365-10370.

Wheeler, C. J. and Croteau, R. (1986) *Arch. Biochem. Biophys.*, **248:** 429-434.

Wierenga, R., Noble, M. E. M. Postma, J. P. M., Groendijk, H., Kalk, K. H., Hol, W. G. J., and Opperdoes, F. R. (1991) *Proteins: Struct. Funct. Gen.*, **10:** 33-49.

Witanowski, M. and Webb, G.A. (1972) in *Annual Reports on NMR Spectroscopy*, (Mooney, E.F., Ed.), Academic Press Inc. (London) Ltd. **Vol. 5A.**, pp. 396-457.

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

Wyckoff, H. W., Hirs, C. H. W., and Timasheff, S. N. (1985a) *Meth. Enzymol.*, Academic Press, **114.**

Wyckoff, H. W., Hirs, C. H. W., and Timasheff, S. N. (1985b) *Meth. Enzymol.*, Academic Press, **115.**

Xu, Z-R., Affleck, R., Wangikar, P., Suzawa, V., Dordick, J.S., and Clark, D.S. (1994) *Biotechnol. Bioeng.*, **43:** 515-520.

Yang, Z. and Robb, D.A. (1993) *Enzyme Microb. Technol.*, **15:** 1030-1036.

Yennawar, N.H., Yennawar, H.P., and Farber, G.K. (1994) *Biochemistry*, **33:** 7326-7336.

Yennawar, N.H., Yennawar, H.P., and Farber, G.K. (1995) *J. Am. Chem. Soc.*, **117:** 577-585.

Young, N. M. and Kaplan, H. (1989) In *Protein Function* (Creighton, T. E., Ed.), IRL Press, Oxford.

Zaks, A. and Klibanov, A. M. (1984) *Science*, **224:** 1249-1251.

Zaks, A. and Klibanov, A.M. (1985) *Proc. Natl. Acad. Sci. U.S.A.*, **82:** 3192-3196.

Zaks, A. and Klibanov, A. M. (1986) *J. Am. Chem. Soc.*, **108:** 2767-2768.

Zaks, A. and Klibanov, A. M. (1988a) *J. Biol. Chem.*, **263:** 3194-3201.

Zaks, A. and Klibanov, A. M. (1988b) *J. Biol. Chem.*, **263:** 8017-8021.

Zhang, M., Huque, E. and Vogel, H.J. (1994) *J. Biol. Chem.*, **269(7):** 5099-5105.

Chapter 2

Use of the pH Memory Effect in Lyophilized Proteins to Achieve Preferential Methylation of α-Amino Groups

1. Introduction

Identification of the amino-terminus of a protein is commonly used as a preliminary indicator of protein identity and homogeneity. Tests for characterizing protein integrity are especially important in the biotechnology industry, where new biotechnological products such as recombinant proteins intended for therapeutic purposes must be of the utmost purity. A common problem encountered in determining the purity of protein preparations is detecting the presence of contaminating proteins of similar size or heterogeneity due to proteolysis during the isolation process. In some instances, such as chymotrypsinogen (Boyer, 1971) and proinsulin (Blundell *et al.*, 1972), proteins are proteolytically processed to give products with two or more polypeptide chains. In each of these cases, multiple α-amino-termini are generated which complicates the characterization of the protein. Moreover, preproteins that are nonhomogeneously proteolyzed to the mature protein may have N- and C-terminal heterogeneities of primary structure. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (Hames, 1990), the commonly accepted technique for the characterization of the global integrity and purity of protein, is unable to distinguish the subtle differences in heterogeneous proteins of similar size.

Conventional N-terminal analysis by identification of acid stable α-amino derivatives (i.e., Edman degradation; Creighton, 1993b) is laborious to implement, especially if such analyses are not carried out routinely. Automated sequencers have

decreased the time and labour involved in determining the amino-terminal sequence of a protein and can be used to determine the presence of multiple free α-amino groups, however they are not readily available in many research institutions and are not a cost-effective way of carrying out simple amino-terminal analysis. Most facile analytical procedures use the instrumental measurement of an intrinsic physical property of the protein or a unique physical property that can be introduced by chemical modification. The usual chemical methods for identifying N-terminal residues require the terminal amino group to have normal chemical properties and to not be blocked covalently. The most common procedures employ fluoro-2,4-dinitrobenzene or dansyl chloride, which form acid-stable linkages with the α-amino group of the N-terminal residue. All amino groups are labeled, the peptide is hydrolyzed to its constituent amino acids, and the amino acid that is labeled on the α-amino group is detected by the yellow colour in the fluorodinitrobenzene procedure or by fluorescence when the dansyl derivative is detected. The labeled N-terminal amino acids can be identified by chromatographic techniques in which each amino acid derivative has a characteristic mobility. A related procedure involves reaction of the α-amino group with cyanate. Acid hydrolysis converts the modified N-terminal residue to the hydantoin. The advantage of the chemical methods of amino-terminal identification is that they do not require highly specialized equipment or instrumentation, but they are time consuming and laborious. Apart from the chemical methods, there is no facile physical technique at present that can be used to identify the amino-terminus of a protein or to determine the presence of multiple amino-termini in a protein preparation.

Nuclear magnetic resonance (NMR) spectrometers are widely available but have not been used for routine protein characterization because of the relatively large amounts of protein required. Recent developments in NMR instrumentation have greatly enhanced its sensitivity, in particular the use of higher field magnets and a DEPT pulse sequence (Bendall *et al.*, 1982). Taralp and Kaplan (1997) have recently demonstrated the feasibility and utility of ^{13}C-iodomethane as a sensitive NMR probe for studying the chemical reactivity of functional groups in proteins under nonaqueous conditions. A feature of this nonaqueous modification procedure was that α- and ε-amino groups were converted to the ^{13}C-trimethylammonium derivatives, which greatly enhanced the sensitivity of detection of individual groups by NMR spectroscopy.

In the present investigation, the nonaqueous methylation procedure was combined with the current NMR technology to yield a sensitive method for analyzing the presence of α-amino groups in protein preparations. Acquisition of a spectrum with the DEPT-135 pulse sequence for four hours yielded a higher quality spectrum than an overnight acquisition with normal ^{13}C-NMR spectroscopy; however, typical acquisition times were 30-40 minutes. The procedure allows for the identification of the number and types of N-terminal derivatives and provides a facile method to investigate the homogeneity of a protein. Its potential use as an analytical tool is demonstrated.

The value of the chemical shift of the various ^{13}C-trimethylammonium derivatives depends on the identity of the adjacent amino acid residue in the protein (Taralp and Kaplan, 1997), and hence, its direct assignment in the intact protein can only be tentative. The chemical shift assignments are based on the resonances of the trimethylated α-amino groups in standard amino acids and di- or tripeptides, which can be appropriately chosen

for well-characterized proteins. However, in unknown or poorly characterized proteins, it is difficult to predict what choice of standard peptide is appropriate to identify the N-terminus, as both the unknown N-terminal and adjacent residues contribute to the observed chemical shift value. Therefore, in addition to characterization of the N-terminal amino acid residue in several intact proteins, the ^{13}C-NMR spectra of ^{13}C-methylated protein hydrolysates were studied. This was feasible because the ^{13}C-trimethylated amino groups of protein are acid stable. The heterogeneity in the protein preparation can then be identified by assignment of its chemical shift value to that of one of the twenty standard ^{13}C-trimethylated amino acids. It is also demonstrated that ^{13}C-methylation of amino groups can be used to verify whether a protein has a free or blocked amino-terminus.

The various classes of functional groups in proteins differ in their ionization constants and there are often differences in the ionization constants of groups within the same class due to their microenvironment in the protein. At a particular pH value, one group could be mostly in the protonated and nonreactive ionization state whereas another group with a lower pK_a value would be in the deprotonated and reactive state. However, as a general rule, it has not been possible to take advantage of this property to achieve selective modification among functional groups of a protein, in particular groups of the same class, by controlling the pH at which the chemical modification is carried out. This is due to the fact that the dynamic protonic equilibria in aqueous solution and the operation of Le Chatelier's principle lead to extensive modification of groups even when only a small fraction of an ionizable group is present in the reactive deprotonated form.

It has been demonstrated that lyophilized enzymes have a pH memory, that is, the activity of the lyophilized enzyme in organic solvents parallels its pH-activity profile of the aqueous solution from which it was prepared (Zaks and Klibanov, 1988). This result implies that the fraction of an ionizable functional group in the deprotonated, and hence reactive form, in a lyophilized protein is determined by the pH of the solution from which it was prepared. As there are no dynamic protonic equilibria in lyophilized proteins, only the nucleophilic deprotonated form will react. None of the protonated form will be converted to the reactive deprotonated form since Le Châtelier's principle does not apply in the nonaqueous environment.

In the present study, the use of the pH memory effect in lyophilized proteins to achieve preferential methylation of α-amino groups is demonstrated. α-Amino groups have expected pK_a values in the range of 6.8 - 8.0 and ε-amino groups from 10.4-11.1 (Creighton, 1993a). Advantage is taken of the difference in ionization constants to achieve methylation of α-amino groups without substantial modification of the ε-amino groups. As will be discussed in detail in Chapter 3, protein carboxyl groups have been observed to react with ^{13}C-iodomethane to form carboxyl ^{13}C-methyl esters whose resonances lay in the same region of the NMR spectrum as those of the ^{13}C-trimethylated amino acid residues. To clarify the spectra, the resonances of these groups were removed by treatment of the samples with hydroxylamine (a carboxyl methyl ester-specific reagent (Jencks, 1958)) or at alkaline pH and with heat, leaving only the resonances of the ^{13}C-trimethylated amino groups.

2. Materials and Methods

2.1 Proteins

Aprotinin, cytochrome c (horse heart, Type III; Lot 62F-7415), ovalbumin (grade V; Lot 106H7070), lysozyme (chicken egg-white, grade I; Lot 89F8275), ribonuclease A (bovine; Lot 127H1104), insulin (bovine; Lot 88H1307; protein powder), α-chymotrypsin (bovine; Lot 91H7195), chymotrypsinogen A (bovine pancreas; Lot 39F8095; Type II), and trypsin inhibitor (soybean, Type II-S; Lot 34H7050) were purchased from Sigma Chemical Company. Albumin (human; Lot M66206; Albuminar-25) was purchased from Armour Pharmaceutical Company. High purity insulin and mutant recombinant insulin {[Lys(B28), Pro(B29)]; monomer, no hexamer} solutions intended for human injection were provided by Dr. Harold Rode of the Biologics and Genetic Therapies Directorate, Health Canada, and were either of bovine or pork origin from licensed manufacturers. *Bacillus thuringiensis* toxin was prepared from crystal protein via the published procedure of Bietlot *et al.* (1989).

2.2 Chemicals and Solvents

^{13}C-Iodomethane was purchased from Sigma-Aldrich Chemical Company. Amino acids, poly-L-Lys·HBr, PheGlyGly, and N^{ε}-Ac-L-Lys were purchased from Sigma-Aldrich Chemical Company. All other chemical, reagents, and solvents were high-purity preparations obtained from commercial sources.

2.3 Preparation of ^{13}C-Methylated Amino Acid or Peptide Standards

Amino acid or poly-L-Lys·HBr or PheGlyGly (1-2 mg) was dissolved in distilled water, the pH of the solution was adjusted to 10.5 with 5 N NaOH, and the sample was lyophilized from a 2 mL volume in a Pyrex screw-cap tube. Amino acids that did not react fully to their trimethylamino derivatives were prepared by an alternative protocol in

order to prevent the amino acids from sticking to the Pyrex tube in semi-crystalline form, which was observed to occur with the original protocol. To amino acid or PheGlyGly peptide (2 mg; e.g., 200 μL of a stock solution of amino acid precursor containing 12 mg in 1200 μL distilled water) was added 300 μL distilled water and the sample was adjusted to pH 10.5 with base, and freeze-dried from a volume of 500 μL in a 0.6 mL eppendorf tube; the latter was placed in a large Pyrex hydrolysis tube. Methylation was carried out *in vacuo* at 100°C for 4 days with ^{13}C-iodomethane (10 μL for each amino acid or 30 μL for poly-L-Lys·HBr) according to the procedure described by Taralp and Kaplan (1997) for proteins. The solution "pH" or pD was adjusted to 11.9 with 5 N NaOH and the sample was heated at 75°C for 30 minutes, followed by a readjustment in the solution pD to 2.5 with 1 N and 6 N HCl. Samples were analyzed by ^{13}C-NMR spectroscopy as indicated below.

N^{α}-trimethylamino Lys standard was synthesized from N^{ε}-Ac-L-Lys precursor. The latter was methylated in the same manner as described above for amino acids. The N^{α}-trimethylamino N^{ε}-Ac-L-Lys was hydrolyzed *in vacuo* for 24 h in 6 N HCl at 110°C to yield the N^{α}-^{13}C-trimethylamino Lys standard.

N^{α}-^{13}C-Trimethylamino PheGlyGly was hydrolyzed *in vacuo* for 1, 3 and 8 days in 6 N HCl at 110°C, and each ^{13}C NMR spectrum was compared to that for the intact methylated tripeptide derivative.

2.4 *In Vacuo* Methylation of Proteins at LpH 6, LpH 7 and LpH 10

Proteins (20 mg) were lyophilized (without buffer, unless otherwise indicated) from the following volumes after adjustment of the solution pH: soluble protein (1 mL); insulin protein powder (40 mL); chicken egg-white lysozyme, LpH 10 (5 mL). Samples

were flash frozen with liquid nitrogen to limit proteolysis. *In vacuo* methylations with ^{13}C-iodomethane (25 μL) were carried out in sealed reaction vessels that were placed in an oven at 75°C for 24 h according to the procedure described by Taralp and Kaplan (1997). After the reaction, excess reagent was recovered by trapping with liquid nitrogen as previously described (Taralp and Kaplan, 1997). Reaction vessels were cracked open and dried by extensive evacuation (typically 2-3 h) to remove residual reagent and ^{13}C-methanol side-product.

Proteins in solution (e.g., insulin and recombinant mutant human insulin samples intended for injection (36 mg per vial, 16-18 mg per reaction); human albumin (Albuminar-25; 300 mg stock, 20 mg per reaction)) were dialyzed (against 2-3x 4 L distilled water using 3500 (insulins) or 12-14000 (human albumin) MWCO dialysis tubing) to remove excipients, lyophilized from larger volumes (20-40 mL), and transferred to reaction vessels in preparation for the *in vacuo* methylation reactions as described above. Methylated insulins or mutant insulins were analyzed by ^{13}C-NMR spectroscopy in their entirety or split equally by mass. If the samples were split, one-half of each of the methylated proteins (8-10 mg) was acid hydrolyzed. One sample of insulin was acid hydrolyzed for 1, 3, and 8 days (to illustrate that trimethylamino PheVal is very difficult to hydrolyze due to steric hindrance). Unhydrolyzed and hydrolyzed methylated proteins were dissolved in urea/D_2O and D_2O, respectively, for ^{13}C-NMR spectroscopic analysis.

During the initial stages of our investigation, samples were prepared with buffer (200 μL 200 mM sodium phosphate, pH 6 or 7 or 200 μL 200 mM sodium metaborate, pH 10) and adjusted to the appropriate pH with acid or base. However, once it was

discovered that methyl esters were being formed under the *in vacuo* methylation conditions, it was observed that phosphate and metaborate buffers also became methylated or esterified, which resulted in the observation of extra peaks in the region of interest in the ^{13}C-NMR spectra. Consequently, buffer was not used; however, some samples of unhydrolyzed and hydrolyzed ^{13}C-methylated proteins presented in the results were prepared with buffer. It is important to note that the buffers enhanced the reactivity of the amino groups, especially of the ε–amino groups at LpH 10. Moreover, the use of buffers illustrated the striking selectivity of the amino groups relative to unbuffered samples (see Chapter 3).

Bt toxin (73 mg) was prepared with 200 µL 0.2 M sodium metaborate buffer, pH 10 and enough 1 N NaOH to adjust the solution pH to 10. The sample was lyophilized (LpH 10) and reacted *in vacuo* with ^{13}C-iodomethane as described above. After reaction, as much sample as possible was dissolved in distilled water and insolubles were separated out using a microfuge; the supernatant was analyzed by ^{13}C-NMR spectroscopy (Varian 300 MHz spectrometer). After the first NMR, the sample was dialyzed against 8 M urea and then against distilled water, lyophilized and re-analyzed by ^{13}C-NMR spectroscopy (prepared with 700 µL urea/D_2O, 50 µL phosphate buffer, pH 7.5, and 10 µL acetonitrile). The sample was subsequently dialyzed against distilled water after the second NMR spectrum was acquired, lyophilized in a hydrolysis tube and hydrolyzed with 6 N HCl, and prepared for a third NMR analysis (Bruker 500 MHz spectrometer) with 800 µL D_2O, 30 µL acetonitrile adjusted to pD 7.9 with NaOH. The spectra of the unhydrolyzed (first NMR) and the hydrolyzed (third NMR) samples

are shown in this chapter; the spectrum of the unhydrolyzed methylated *Bt* toxin that was dialyzed against urea (second NMR) is shown in Chapter 3 (Section 2.4).

2.5 Reaction of ^{13}C-Methylated Protein with Hydroxylamine and Sodium Hydroxide

After initial ^{13}C-NMR spectroscopic analysis, the ^{13}C-methylated protein was reacted with 2.5% hydroxylamine at pH 9 (Darbre, 1986a; Halpin and Richardson, 1985). Samples were heated at 78°C for 30 min in screw-cap NMR tubes (5 mm; Wilmad Glass Co.), then either analyzed directly or frozen at -20°C until they were analyzed via ^{13}C-NMR spectroscopy. If necessary, samples were treated further with base, by adjusting to pD 11.9 with 1 N NaOH, and heated for an additional 30 min at 78°C, followed by a readjustment of the pD to 7.9 with 1 N HCl, and subsequently analyzed by ^{13}C-NMR spectroscopy.

2.5.1 Aqueous Reaction of ^{13}C-Methylated Ribonuclease A with Hydroxylamine and Sodium Hydroxide: A ^{13}C-NMR Spectroscopic Kinetic Study for the Identification of ^{13}C-Resonances Arising from ^{13}C-Methyl Esters

^{13}C-Methylated ribonuclease A (LpH 5, 48 h, 75°C) was analyzed by ^{13}C-NMR spectroscopy in 8 M urea/D_2O in the presence of acetonitrile (4 spectra acquired). After acquisition of the initial spectrum (spectra), the sample (approximately 800 μL; see Section 2.8) was reacted with 2.5% hydroxylamine (30 μL 50 wt.% hydroxylamine in water, pH 9) at room temperature (9 spectra acquired at pD 8.7) and the disappearance of the resonances corresponding to the ^{13}C-methyl esters of protein carboxyl groups was observed at specific time blocks over the course of a multi-hour acquisition. The sample was stored at −10°C overnight in between acquisitions when required. The pD of the aqueous reaction with hydroxylamine was initially at 8.7, and then increased to pD 10.0 (5 spectra acquired). The reaction was carried out at room temperature for 7.5 h and the

sample was heated at 75°C for two 30 min intervals and then for 2 h (3 spectra acquired). In order to find the end-point of the experiment, the pD of the solution was increased to 12.0 and the sample was heated for 1 h at 75°C. The sample was then dialyzed against 3x 4 L distilled water (3500 MWCO dialysis tubing), speed vacuumed, and a final spectrum was acquired at pD 5.4. Another sample of ^{13}C-methylated ribonuclease A (LpH 5, 72 h, 75°C) was used to study the hydroxylamine reaction. The sample pD was adjusted to 9.9 after the addition of hydroxylamine (30 μL 50 wt.% hydroxylamine in water, pH 9; final concentration was 2.5%) and the sample was heated at 75°C for two 30 min intervals.

The ^{13}C-NMR spectroscopic kinetic study was carried out with ^{13}C-methylated ribonuclease A (LpH 5, 48 h, 75°C) for the identification of the ^{13}C-resonances that arise from ^{13}C-methyl esters. Removal of the ^{13}C-methyl esters arising from Asp and Glu residues (53.46 and 53.18 ppm, respectively; see Chapter 3) was slow in the presence of hydroxylamine at pD 8.7; there was a slow but progressive increase in the resonance corresponding to ^{13}C-methanol. An increase in the solution pD to 10.0 increased the conversion of methyl esters to the hydroxamate derivatives; however, the resonance for the carboxyl methyl esters of Asp residues did not diminish completely (after 6 h in the presence of hydroxylamine at room temperature), even after the sample was heated for a total of 3 h. (The resonance for the methyl ester of the Asp residues could be qualitatively observed more easily than that of the Glu residues because the latter overlap the resonance for the trimethyl α-amino Lys.) The Asp methyl ester only diminished significantly when the sample pD was increased to 12.0 and the sample was heated for 1 h. The sample was dialyzed and the resonances remaining were those for the trimethyl

ε-amino Lys, trimethyl α-amino Lys and the trimethylamino group resonance derived from an impurity at 54.15 ppm. A separate sample of ^{13}C-methylated ribonuclease A (LpH 5, 72 h, 75°C) was used to study the hydroxylamine reaction. The sample pD was adjusted to 9.9 after the addition of hydroxylamine and the sample was heated at 75°C for two 30 min intervals. The resonance for the methyl esters of the Asp residues decreased substantially, but again, the result was an incomplete conversion to the hydroxamate groups. Thus, it was still necessary to increase the pD to 12.0 and heat for an additional hour or longer in some cases. These results led to the trial of an *in vacuo* hydroxylamine reaction (see Section 2.5.2).

2.5.2 *In Vacuo* Reaction of ^{13}C-Methylated Protein with Hydroxylamine to Form Hydroxamates with Carboxyl ^{13}C-Methyl Esters and ^{13}C-Methanol

For some proteins (e.g., α-chymotrypsin, ovalbumin, human albumin), the hydroxylamine reaction in aqueous medium resulted in a gel state, which is indicative of an intermediate denaturation state (Magdassi, 1996; Ferry, 1948). For these samples (in screw-cap NMR tubes), complete conversion of methyl esters to hydroxamates for the purpose of identifying the ^{13}C-resonances that arose from ^{13}C-methyl esters in the region of the ^{13}C-NMR spectra between 50-57 ppm was not possible. The elimination of the ^{13}C-methyl ester peaks clarified the spectra and allowed for the identification of N-terminal resonances. The resultant gel state also led to the need for overnight NMR acquisitions in order to obtain a good quality spectrum with a good signal-to-noise ratio. In order to circumvent this problem, the hydroxylamine reaction was carried out *in vacuo* on the dried derivatized protein [^{13}C-methylated ribonuclease A (10-20 mg)] at 75°C with hydroxylamine (30 μL 50 wt.% hydroxylamine in water, pH 9) at various times (0, 15, 30 min, 1, 2, 4, 8, 12, and 24 h). After reaction for up to 2 h the protein was in an

insoluble gel state, and thus, inconvenient form for ^{13}C-NMR spectroscopic analysis (line-broadening problem). The 4 h time point gave the optimum result with respect to protein solubility. It was sufficient time to remove all resonances due to ^{13}C-methyl esters from the spectra of ^{13}C-methylated ribonuclease A without significant degradation of the protein, which occurred with longer reaction times. Thus, although it took longer (four hours at 75°C) to achieve complete hydroxamate derivatization of carboxyl methyl esters compared to the aqueous protocol, the *in vacuo* procedure was found to be better than the latter for our purposes because it was possible to solubilize the problematic proteins. The latter bypassed the gel state denaturation phase, which also was evident in the *in vacuo* protocol after 2 h of heating. The 4 h of heating is the optimal time in terms of solubility and also in terms of exhaustive degradation of observable trimethylamino resonances.

2.6 Acid Hydrolysis of ^{13}C-Methylated Protein

^{13}C-Methylated protein was transferred to a Pyrex hydrolysis tube and hydrolyzed in 6 N HCl *in vacuo* at 110 °C for 24 h (Darbre, 1986b). One sample of insulin was acid hydrolyzed for 1, 3, and 8 days (to illustrate that trimethylamino PheVal is very difficult to hydrolyze due to steric hindrance). Samples were dried in a desiccator under vacuum with phosphorus pentoxide (P_2O_5) and sodium hydroxide pellets, and dried a second time following the addition of 1 mL distilled water to remove residual acid.

α-Chymotrypsin and chymotrypsinogen A samples were acid hydrolyzed (Darbre, 1986b) and subsequently performic acid oxidized (Hirs, 1956) prior to ^{13}C-NMR spectroscopic analysis.

2.7 Performic Acid Oxidation of Methylated α-Chymotrypsin and Chymotrypsinogen Hydrolysates

Performic acid oxidation of α-chymotrypsin and chymotrypsinogen A samples was carried out according to the procedure of Hirs (1956; referenced in Bailey, 1964). Performic acid solution was prepared by incubating a mixture of 0.5 mL 30% hydrogen peroxide and 9.5 mL of 98-100% formic acid at room temperature for 2 h in a screw cap hydrolysis tube. The mixture was divided equally and is sufficient to react two 30 mg samples of protein. Normally, performic acid oxidation is carried out on protein, but it was performed here on hydrolyzed protein, which should not pose a problem. Each protein hydrolysate sample was dissolved in 2.5 mL of 99% formic acid and 1.0 mL anhydrous methanol in a Pyrex hydrolysis tube; the latter is added to prevent freezing of the solution later on in the protocol. The three hydrolysis tubes containing the two protein hydrolysate solutions and the performic acid solution were placed in an ice-NaCl bath at −10°C and cooled for 30 min, at which time the performic acid solution was split and added to each sample. After a reaction time of 2.5 h, the content of each tube was rinsed with 50 mL of ice water and transferred into a 1 L round bottom flask containing 300 mL of water at 0°C. The samples were shelled with liquid nitrogen and lyophilized for 3 days. Samples were re-dissolved in 10 mL water and re-lyophilized (in a scintillation vial or pointy bottom flask for facile sample retrieval) to remove traces of performic acid. Samples were fluffy light beige lyophilized powders.

2.8 NMR Spectroscopic Analysis of ^{13}C-Methylated Proteins and Protein Hydrolysates

^{13}C-NMR spectra were obtained using a Bruker spectrometer operating at 9.4 Tesla (^{1}H, 400.13 MHz; ^{13}C, 100.6 MHz). ^{13}C-Methylated protein was dissolved in deuterium oxide containing 8 M ultra pure urea (700 μL; pH 3.8; Schwarz/Mann Biotech;

Lot 4116); hydrolyzed ^{13}C-methylated protein was dissolved in deuterium oxide. Acetonitrile (30 μL) was added to the protein sample as an internal chemical shift reference (Breitmaier, E., and Voelter, W., 1987a) ($^{13}CH_3$, +1.70 ppm). For quantification, 3-pentanol (40 μL) ($^{13}CH_3$, 10.14 ppm; $^{13}CH_2$, 29.44 ppm; ^{13}CH, 75.41 ppm) was included as internal standard (Breitmaier, E., and Voelter, W., 1987b). The sample solution was adjusted to pD 5.4 (to preserve methyl esters) using 0.1 or 1 N HCl and/or NaOH, transferred to an NMR tube (5 mm), and analyzed. Protein hydrolysate (and some unhydrolyzed methylated proteins; prepared in this manner prior to the discovery that carboxyl groups become esterified by the *in vacuo* methylation procedure) solutions were adjusted to pD 7.9 with 70 μL 1 M sodium phosphate buffer, pH 7.5 and acid or base if required. Spectra were acquired typically for 600 scans using the DEPT-135 pulse sequence (Bendall *et al.*, 1982). For the hydroxylamine and base treated samples, the ^{13}C-NMR spectra were acquired at pD 9.4 and 7.9, respectively. It is noteworthy that for hydroxylamine and base treated α-chymotrypsin, the spectra were acquired overnight (18000 scans) to improve the signal-to-noise ratio (Figure 2-3c), as the protein was observed to be in a gel state after these treatments. The latter problem was circumvented by carrying out the hydroxylamine reaction *in vacuo* for 4 h at 75°C (see Section 2.5.2, Chapter 2; Figure 2-3d).

2.8.1 Choice of Internal Standard: ^{13}C-Hexamethylethylenediamine·2HI vs 3-Pentanol

^{13}C-Hexamethylethylenediamine·2HI (^{13}C-HMED·2HI) [($^{13}CH_3$ = 29.1% ^{13}C-enriched), +54.6 ppm] was synthesized (see below for synthetic protocol) and employed as an internal standard. This compound was useful for qualitatively comparing samples reacted at LpH 6, 7, and 10, since its chemical shift was within the region of

interest. However, it became necessary to find and employ another compound, namely 3-pentanol, as an internal standard in samples which were treated with hydroxylamine and base, since the ^{13}C-HMED·2HI standard degraded under such treatment in the presence of 9 M urea. However, ^{13}C-HMED·2HI is reasonably easy to make and is an excellent internal standard if further manipulation of the sample is not performed; for NMR analysis of a 20 mg protein sample, use 10 μL of 50 mg ^{13}C-HMED·2HI dissolved in 200 μL (250 μL total volume; 0.5 M ^{13}C-HMED·2HI; approximately 2 mg per sample) stock in D_2O. 3-Pentanol has a high boiling point and apparently does not degrade, however, it was observed for some methylated protein samples treated with hydroxylamine that the peaks of interest appeared larger than they should have been when the internal standard peak was set to 10 cm height upon data processing; the baseline also increased. In such cases the data was processed such that the N^{ε}-trimethylamino Lys resonance was kept at a constant height, which is a sensible assumption. It was possible to make a reasonable comparison between spectra acquired before and after hydroxylamine or base treatments.

Acetonitrile was not chosen as an internal standard because it is volatile and was observed to react with hydroxylamine and base to form a derivative, which may involve a proton transfer [CH_3-C(=NH-NH-OH)], with a characteristic resonance at 16.0 ppm. However, acetonitrile was included in all samples as a chemical shift reference compound. Isopropanol was also tested as a potential internal standard, but was observed to create a new resonance at 13.65 ppm. Besides the chosen 3-pentanol internal standard, glycerol ($^{13}CH_2$, -63.56 ppm; ^{13}CH, 73.18 ppm; 100 μL 5.0 M glycerol or 501 μmol) was a good and cheap alternative, but was not employed because it gives a negative resonance

at -63.56 ppm, which was esthetically unappealing if it were desired to illustrate a spectrum for publication of the chemical shift region of interest between 47-57 ppm and the positive resonance of the glycerol internal standard at 73.18 ppm.

2.8.2 Synthesis of ^{13}C-Hexamethylethylenediamine·2HI Internal Standard

^{13}C-Hexamethylethylenediamine was synthesized *in vacuo* from tetramethylethylenediamine (250 μL) and ^{13}C-iodomethane (250 μL) in the presence of octane (250 μL). The tetramethylethylenediamine was placed in a Pyrex tube and frozen in liquid nitrogen. The tube was narrowed and ^{13}C-iodomethane was added to the reaction vessel and the tube was sealed under vacuum. Upon warming of the sample to room temperature, the reaction proved to be highly spontaneous (caution!), as a white salt formed and splashed to cover the entire surface area of the reaction vessel. This could have been prevented by using more octane as solvent to stabilize the mixture, as observed for the same reaction carried out with a huge excess of unlabeled iodomethane and 1 mL octane. Sealed tubes were cracked opened and samples were dried under vacuum in a desiccator. The ^{13}C-NMR spectrum of the labeled product revealed that the reaction had gone to approximately 85% completion relative to the initial tetramethyl-ethylenediamine. Approximately half the sample (372.28 mg) was reacted to completion in an immiscible biphasic system containing iodomethane (2 mL) and distilled water (0.5 mL) in a scintillation vial overnight (16 h) at room temperature with stirring. After the reaction, the sample was heated on low to evaporate the water and excess reagent, followed by evacuation under vacuum prior to analysis by ^{13}C-NMR spectroscopy (single peak observed at 54.6 ppm). Since the ^{13}C-NMR spectrum revealed that the reaction had

gone to 85% completion with ^{13}C-iodomethane, the overall enrichment with carbon-13 atoms was 29.1 % in the ^{13}C-HMED·2HI internal standard compound.

2.9 NMR Spectroscopic Analysis of ^{13}C-Methylated Amino Acid Standards

^{13}C-NMR spectra were obtained using a Bruker spectrometer operating at 4.7 Tesla (^{1}H, 200.13 MHz; ^{13}C, 50.32 MHz). ^{13}C-Methylated amino acid was dissolved in deuterium oxide (400 μL), followed by the addition of acetonitrile (25 μL), which was included as an internal chemical shift reference (Breitmaier and Voelter, 1987a). Sample "pH" or pD [pD = pH meter reading + 0.4; Glasoe and Long, 1960] was adjusted to 2.8 with 1 N HCl and/or 1 N NaOH with vortex mixing. If required, samples were centrifuged using a microcentrifuge at 15000 g for 2 minutes to remove insoluble material. The supernatant was transferred to an NMR tube (5 mm), and analyzed. ^{13}C-Methylated poly-L-Lys·HBr was first acid hydrolyzed and dried prior to ^{13}C NMR spectroscopic analysis. Spectra were acquired typically for 400 scans using the DEPT-135 pulse sequence (Bendall *et al*, 1982).

Betaine (20 mg; trimethylamino Gly; natural abundance carbon-13) was analyzed by ^{13}C-NMR spectroscopy at acidic and neutral pD values (2.8, 7.6) in the presence (B) and absence (NB) of buffer (pH 2.1 buffer containing acetic and formic acids; pH 7.5 buffer containing sodium phosphate) in order to determine whether there is a large change in the chemical shift of the ^{13}C-methylated amino acid derivatives with pD. It was necessary to establish the latter if comparisons were to be made with the work of Taralp (Doctoral Thesis, 1997), who established a list of chemical shifts for the methylated amino acid derivatives. The chemical shifts observed were similar: pD 2.8,

NB (54.25 ppm); pD 2.8, B (pD 2.5 B) (54.24 ppm); pD 7.6, NB (54.15 ppm); and pD 7.5, B (54.20 ppm).

3. Results

3.1 Aqueous Reaction of ^{13}C-Methylated Ribonuclease A with Hydroxylamine and Sodium Hydroxide: A ^{13}C-NMR Spectroscopic Kinetic Study for the Identification of ^{13}C-Resonances Arising from ^{13}C-Methyl Esters

A ^{13}C-NMR spectroscopic kinetic study was carried out with ^{13}C-methylated ribonuclease A (LpH 5, 48 h, 75°C) for the identification of the ^{13}C-resonances that arise from ^{13}C-methyl esters. Removal of the ^{13}C-methyl esters arising from Asp and Glu residues (53.46 and 53.18 ppm, respectively; see Chapter 3) was slow in the presence of hydroxylamine at pD 8.7; there was a slow but progressive increase in the resonance corresponding to ^{13}C-methanol. An increase in the solution pD to 10.0 increased the conversion of methyl esters to the hydroxamate derivatives; however, the resonance for the carboxyl methyl esters of Asp residues did not diminish completely (after 6 h in the presence of hydroxylamine at room temperature), even after the sample was heated for a total of 3 h at 75°C. (The broadened resonance for the methyl ester of the Asp residues could be qualitatively observed more easily than that of the Glu residues because the latter overlaps the resonance for the trimethyl α-amino Lys.) The Asp methyl ester only diminished significantly when the sample pD was increased to 12.0 and the sample was heated for 1 h. The sample was dialyzed and the resonances remaining were those for the trimethyl ε-amino Lys, trimethyl α-amino Lys and the trimethylamino group resonance derived from an impurity at 54.15 ppm. A separate sample of ^{13}C-methylated ribonuclease A (LpH 5, 72 h, 75°C) was used to study the hydroxylamine reaction. The sample pD was adjusted to 9.9 after the addition of hydroxylamine and the sample was

heated at 75°C for two 30 min intervals. The resonance for the methyl esters of the Asp residues decreased substantially, but again, the result was an incomplete conversion to the hydroxamate groups. Thus, it was still necessary to increase the pD to 11.9 and heat for an additional hour or longer in some cases. These results led to the trial of an *in vacuo* hydroxylamine reaction (see Section 2.5.2).

Thus, in some cases, the removal of the methyl ester resonances from the ^{13}C-NMR spectrum was not complete with either hydroxylamine or base treatment or both in aqueous medium. In other cases, such as was the case with α-chymotrypsin, ovalbumin and human serum albumin, the proteins gelled after hydroxylamine treatment and the conversion of methyl esters to hydroxamates was incomplete. For these reasons, an *in vacuo* hydroxylamine reaction was carried out on ribonuclease A. The removal of the methyl ester resonances was achieved with an optimal 4 h reaction time with respect to protein solubility for the NMR analysis (see Section 2.5.2). Hydroxylamine is able to cleave protein at Asn-Gly peptide bonds as a result of the tendency of the asparaginyl side-chain to cyclize (in the absence of steric hindrance imposed by a side-chain on the next amino acid, Gly) forming a substituted succinimide that is susceptible to nucleophilic attack by hydroxylamine (Darbre, 1986). In the present study of N-terminal identification, structural integrity is not a priority; however, if maintainance of protein structure is desired, a saponification reaction under mild conditions could be carried out to identify and remove the resonances arising from carboxyl methyl esters (Broomfield *et al.*, 1965).

3.2.1 Intact ^{13}C-Methylated Proteins

The results obtained using bovine insulin as a test protein demonstrate that by controlling the pH of lyophilization (LpH) it is possible to methylate the α-amino groups with only a minor amount of methylation of the solitary lysine ε-amino group (Protein Data Bank, ID: 9INS). Figure 2-1 shows the ^{13}C-NMR spectra obtained with ^{13}C-methylated bovine insulin under various conditions. Methylation at LpH 10 gives several peaks in the region from 52 to 57 ppm (Figure 2-1a). On treatment with base or hydroxylamine several of these peaks decreased and eventually disappeared after treatment for 30 to 60 minutes (Figure 2-1b), with the appearance of a peak corresponding to ^{13}C-methanol at 49.8 ppm (Breitmaier, E., and Voelter, W., 1987b). The positions of these resonances (Breitmaier, E., and Voelter, W., 1987c) and their disappearance on treatment with hydroxylamine or base with the concomitant appearance of methanol are consistent with hydroxylaminolysis of methyl esters formed by reaction of iodomethane with carboxyl groups in the lyophilized protein. The remaining major peaks have previously been identified (Taralp and Kaplan, 1997) and correspond to the ^{13}C-trimethylphenylalanine α-amino-terminus (53.4 ppm), the ^{13}C-trimethyl ε-amino group of lysine (53.6 ppm), the ^{13}C-trimethylglycine α-amino-terminus (55.0 ppm), and the phenolic -O$^{13}CH_3$ group on the side-chain of tyrosine (56.0 ppm). Taralp and Kaplan (1997) were unable to make the assignment for the broad resonance at 53.1 ppm (peak 7) or that for the resonances at 53.74 and 53.97 ppm (peaks 3 and 4; more clearly observed at LpH 7, Figure 2-1c). Peak 7 was deduced to correspond to the overlapping resonances of δ-carboxyl ^{13}C-methyl ester groups of Glu residues, while peaks 3 and 4 were deduced to correspond to the α-carboxyl ^{13}C-methyl ester groups of the two C-terminal residues (see Chapter 3 for details). Methylation at LpH 7 (Figure 2-1c), followed by treatment

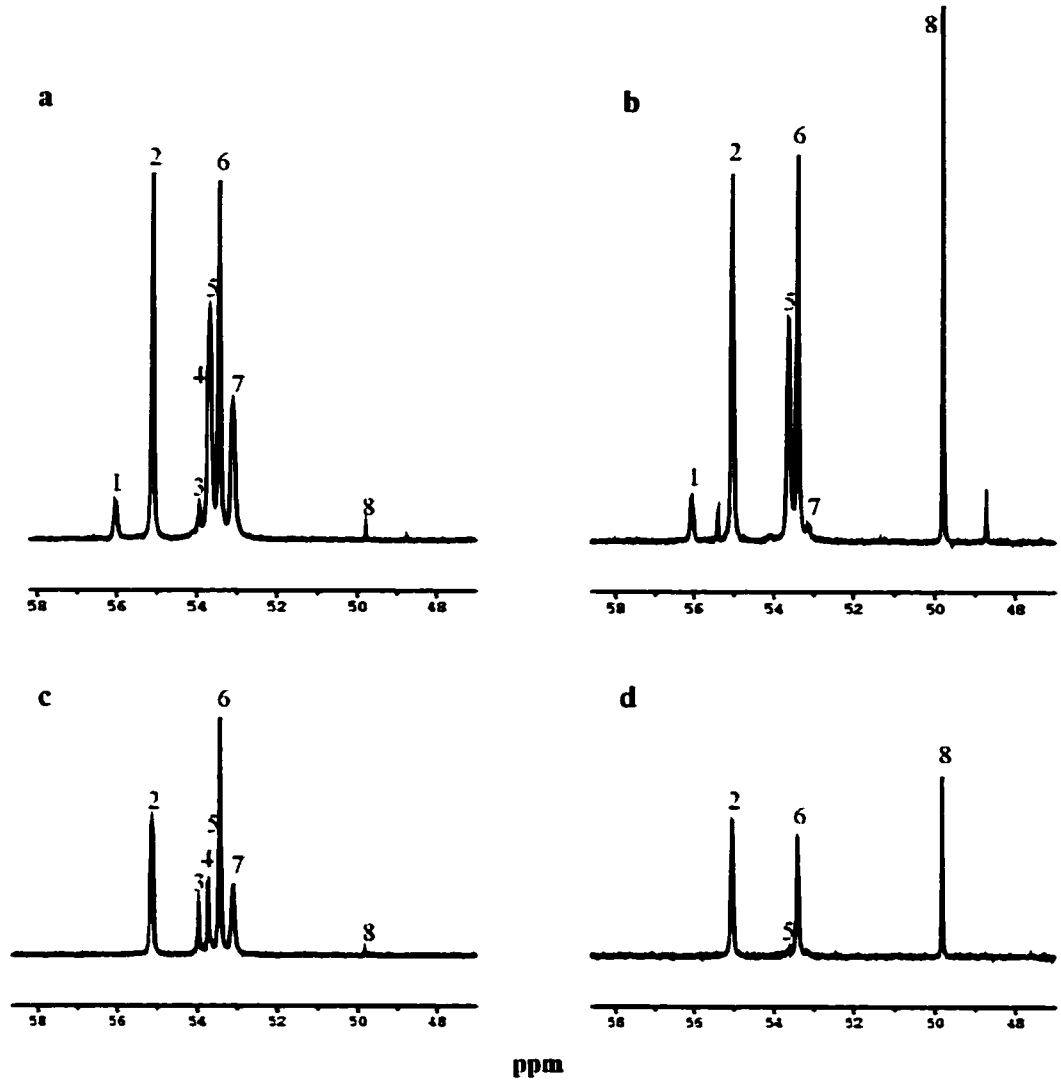


Figure 2-1. Expanded ^{13}C-NMR spectra of ^{13}C-methylated bovine insulin. (a) LpH 10 reaction with ^{13}C-iodomethane (75°C, 24 h); (b) LpH 10 reaction, followed by hydroxylamine, base and heat treatments; (c) LpH 7 reaction; (d) LpH 7 reaction, followed by hydroxylamine, base and heat treatments.

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) Tyr(OMe) (56.04 ppm); (2) $Me_3N^{\alpha+}$-Gly (55.08 ppm); (3) C_αOOMe (53.97 ppm);

(4) C_αOOMe (53.74 ppm); (5) $Me_3N^{\varepsilon+}$-Lys (53.66 ppm); (6) $Me_3N^{\alpha+}$-Phe (53.42 ppm);

(7) C_δOOMe (53.12 ppm); and (8) MeOH (49.79 ppm).

with hydroxylamine and base, gives a spectrum (Figure 2-1d) in which the significant resonances correspond to the two amino-termini and only a trace resonance corresponding to the ^{13}C-trimethyl ε-amino groups of Lys residues is observed. It should be noted that the resonance due to the ^{13}C-trimethylphenylalanine α-amino-terminus (peak 6) decreases in size after removal of the ester resonances, indicating that a carboxyl ester resonance overlapped that of this group (see Chapter 3).

Lysozyme has only one polypeptide chain with lysine at its amino-terminus (Protein Data Bank, ID: 1DPX). The ^{13}C-NMR spectra obtained after methylation of this protein are shown in Figure 2-2. Methylation at LpH 6 (Figure 2-2a) gives a spectrum with three major peaks, one of which (53.5 ppm) disappears after treatment with hydroxylamine (Figure 2-2b), leaving only two major peaks. The resonance at 53.7 ppm corresponds to the ^{13}C-trimethyl ε-amino groups of lysine residues and the resonance at 53.1 ppm is from the ^{13}C-trimethyl α-amino-terminus. Although lysozyme has six ε-amino groups, its single α-amino group gives a larger signal than all six ε-amino groups combined. As in the case of insulin, this observation demonstrates the selectivity of the *in vacuo* methylation reaction for α-amino groups versus ε-amino groups at LpH 6.

Figure 2-3 shows the ^{13}C-NMR spectra obtained for ^{13}C-methylated ribonuclease A and α-chymotrypsin after removal of the methyl ester resonances. Ribonuclease A, like lysozyme has a lysine α-amino-terminus (Protein Data Bank, ID: 5RSA), and its spectrum (Figure 2-3a) is virtually identical to that of lysozyme with only two major resonances at 53.1 ppm and 53.6 ppm. However, a small peak is observable at 54.1 ppm, indicating the presence of a minor protein impurity in the preparation. The

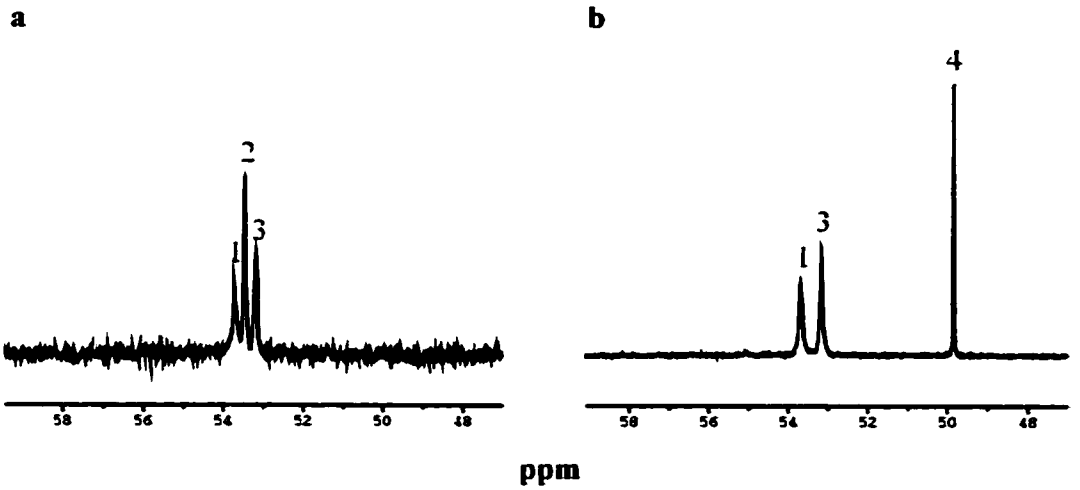


Figure 2-2. Expanded ^{13}C-NMR spectra of ^{13}C-methylated chicken egg-white lysozyme. (a) LpH 6 *in vacuo* reaction with ^{13}C-iodomethane (75°C, 24 h); (b) LpH 6 reaction, followed by hydroxylamine and heat treatments.

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) **Me**$_3$N$^{\varepsilon+}$-Lys (53.73 ppm); (2) C$_\gamma$OO**Me** (53.46 ppm); C$_\delta$OO**Me** (53.18 ppm) overlaps peak (3); (3) **Me**$_3$N$^{\alpha+}$-Lys (53.15 ppm); and (4) **Me**OH (49.83 ppm).

latter resonance was also observed in spectra for α-chymotrypsin and soybean trypsin inhibitor.

α-Chymotrypsin has three amino-termini (Protein Data Bank, ID: 1AB9) and its ^{13}C-NMR spectra before and after the removal of the methyl ester resonances are shown in Figures 2-3b and 2-3c, respectively. The hydroxylamine-treated sample shown in Figure 2-3c revealed a small peak due to the ^{13}C-trimethyl ε-amino groups at 53.6 ppm and at least five resonances from ^{13}C-trimethyl α-amino groups. The resonances from the trimethylated α-amino groups of the A, B and C chains, peaks 6, 7 and 8 respectively, have been identified previously (Taralp and Kaplan, 1997). This shows that the protein preparation contains either other protein impurities or, as is more likely for a proteolytic enzyme, has undergone limited autolysis. Figure 2-3d illustrates the improvement in removing the ester resonances by employing the *in vacuo* hydroxylamine reaction instead of the aqueous reaction for proteins which tend to gel during the latter protocol (see Sections 2.5.2 and 2.7).

Cytochrome c (Protein Data Bank, ID: 1CRC; Dayhoff, 1972) and ovalbumin (Protein Data Bank, ID: 1OVA) have blocked amino-termini. The ^{13}C-NMR spectra obtained for these proteins (Figure 2-4) after appropriate treatments revealed only a single major peak at 53.6 ppm, corresponding to the resonances from the ^{13}C-trimethyl ε-amino groups in these proteins. This result is what would be predicted for a protein with a blocked amino-terminus. In the case of ovalbumin (Figure 2-4a) there is a small but distinct resonance at 52.7 ppm, suggesting the presence of a minor protein impurity with a free α-amino-terminus; however, the peak may represent a residual amount of carboxyl methyl ester due to the fact that the ovalbumin sample gelled with

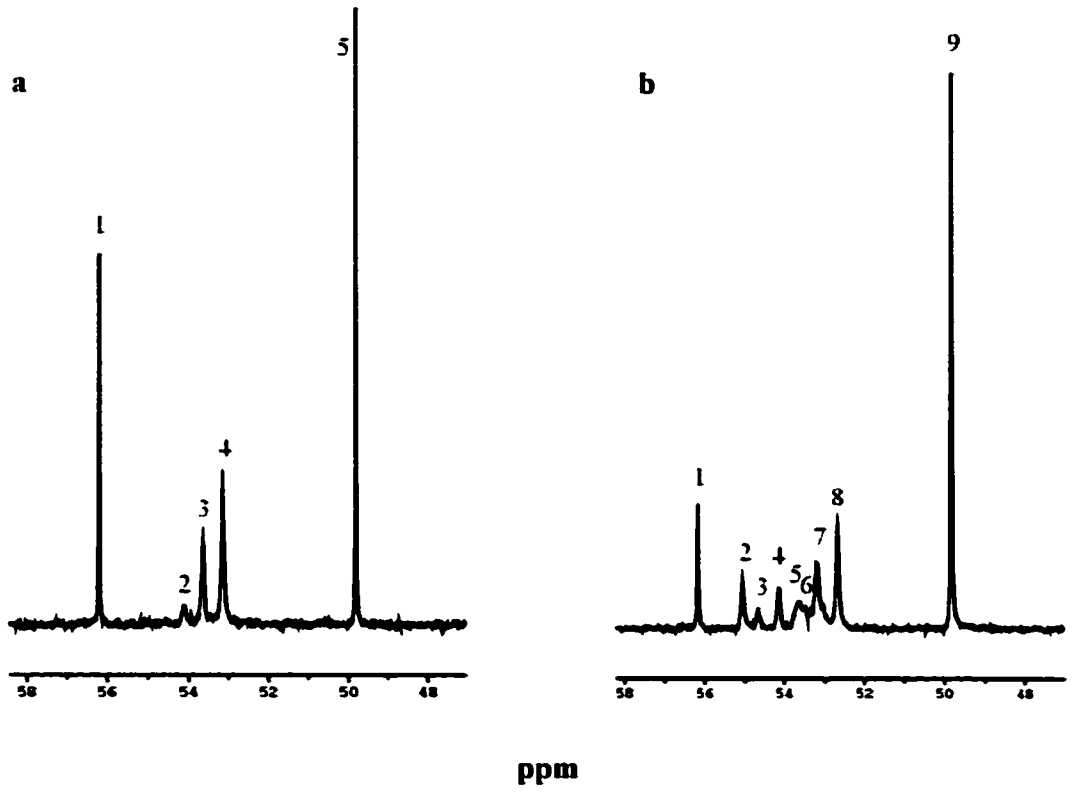


Figure 2-3. Expanded ^{13}C-NMR spectra of ^{13}C-methylated (reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h) (a) ribonuclease A (LpH 6) and (b) α-chymotrypsin (LpH 6) after treatments with hydroxylamine, base and heat in aqueous medium (sample gelled); (c) α-chymotrypsin (LpH 7) before hydroxylamine treatment; (d) α-chymotrypsin (LpH 7) treated *in vacuo* with hydroxylamine for 4 h.

For ribonuclease A (a), peak resonances correspond to the ^{13}C-enriched methyl groups in: (1) unidentified small molecule (56.19 ppm); (2) unidentified α-$^{+}$N**Me$_3$** impurity (54.13 ppm); (3) **Me$_3$**N$^{\varepsilon+}$-Lys (53.63 ppm); (4) **Me$_3$**N$^{\alpha+}$-Lys (53.13 ppm); and (5) **Me**OH (49.81 ppm). The spectrum for untreated (without hydroxylamine) methylated ribonuclease A contained resonances for α-, γ-, and δ-carboxyl ^{13}C-methyl esters at 53.71, (53.49, 53.45), and 53.18 ppm, respectively.

For α-chymotrypsin (b), peak resonances correspond to the ^{13}C-enriched methyl groups in (1) unidentified small molecule (56.17 ppm); (2), (3), (4) unidentified α-$^{+}$N**Me$_3$** (55.16, 54.69, and 54.21 ppm); (5) **Me$_3$**N$^{\varepsilon+}$-Lys (53.68 ppm); (6) **Me$_3$**N$^{\alpha+}$-Cys (53.45 ppm); (7) **Me$_3$**N$^{\alpha+}$-Ile (53.23 ppm); (8) **Me$_3$**N$^{\alpha+}$-Ala (52.71 ppm); and (9) **Me**OH.

Continued...

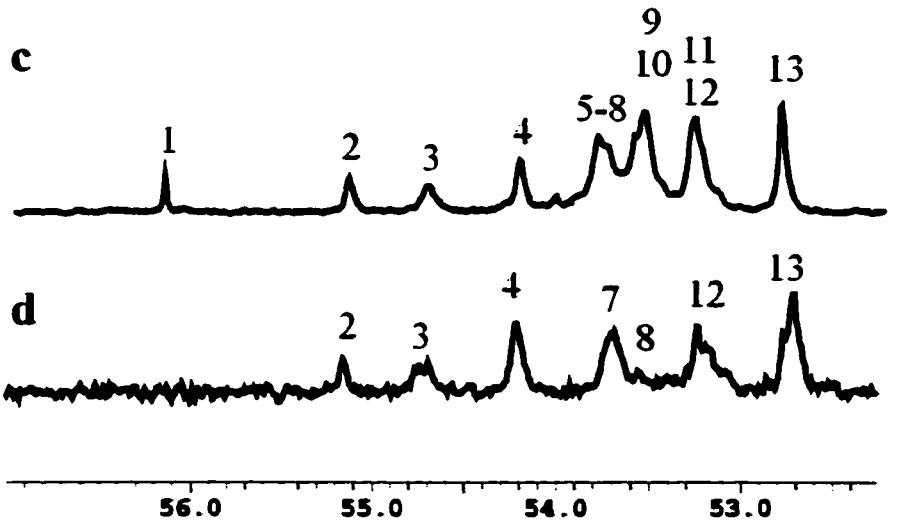


Figure 2-3. For α-chymotrypsin (c) and (d), (1) unidentified small molecule (56.17 ppm); (2), (3), (4) unidentified α-$^{+}$N**Me$_3$**, impurities (55.16 ppm, 54.69 ppm, and 54.21 ppm); (5) C_αOO**Me** (53.95 ppm; C-terminal Asn-245); (6) C_αOO**Me** (53.77 ppm); (7) **Me$_3$**N$^{\varepsilon+}$-Lys (53.68 ppm); (8) C_αOO**Me** (53.72 ppm); (6, 8) C-termini: Leu-13, Tyr-146; (9) C_γOO**Me** (53.50 ppm); (10) half-cystine, Cys-N$^{\alpha+}$**Me$_3$** (53.45 ppm); (11) C_δOO**Me** (53.21 ppm); (12) **Me$_3$**N$^{\alpha+}$-Ile (53.23 ppm); (13) **Me$_3$**N$^{\alpha+}$-Ala (52.71 ppm).

hydroxylamine and base treatments (see Sections 2.5.2 and 2.7). Similarly, there are several very minor impurities with free α-amino groups visible in the spectrum (Figure 2-4b) for the cytochrome c preparation, which are likely residual carboxyl methyl ester resonances (except the peak at 56.0 ppm, which corresponds to the tetramethylammonium ion). Cytochrome c (Protein Data Bank, ID: 1CRC) has 19 lysine residues whereas ovalbumin (Protein Data Bank, ID: 1OVA) has only 3, and the intensity of the resonance of its ^{13}C-trimethyl ε-amino groups was observed to be correspondingly greater (Figure 2-4a and 2-4b). Human albumin has an N-terminal Asp residue and has a similar spectrum to that of ovalbumin (spectrum not shown). The spectrum of methylated human albumin treated with hydroxylamine *in vacuo* revealed a single resonance at 53.7 ppm corresponding to the resonances from the ^{13}C-trimethyl ε-amino groups (Figure 2-4, c and d). (The ^{13}C-methanol resonance was not observed because the sample was evacuated after the reaction.) A resonance for the trimethyl-α-amino group of Asp was not observed.

Soybean trypsin inhibitor (STI) has an Asp N-terminal residue. Figures 2-5a and 2-5b show the spectra for ^{13}C-methylated STI (LpH 7) before and after the removal of the methyl ester resonances. The main peak observed after *in vacuo* hydroxylamine treatment was that of the ^{13}C-trimethyl ε-amino groups at 53.7 ppm. The resonance at 54.2 ppm is a trimethylamino group of a minor protein impurity, while that at 53.3 ppm likely corresponds to the ^{13}C-trimethyl α-amino groups of the Asp N-terminal residue. The latter resonance is very small, as expected, since it was difficult to synthesize and observe (by ^{13}C-NMR spectroscopy) the ^{13}C-trimethylamino Asp standard compound.

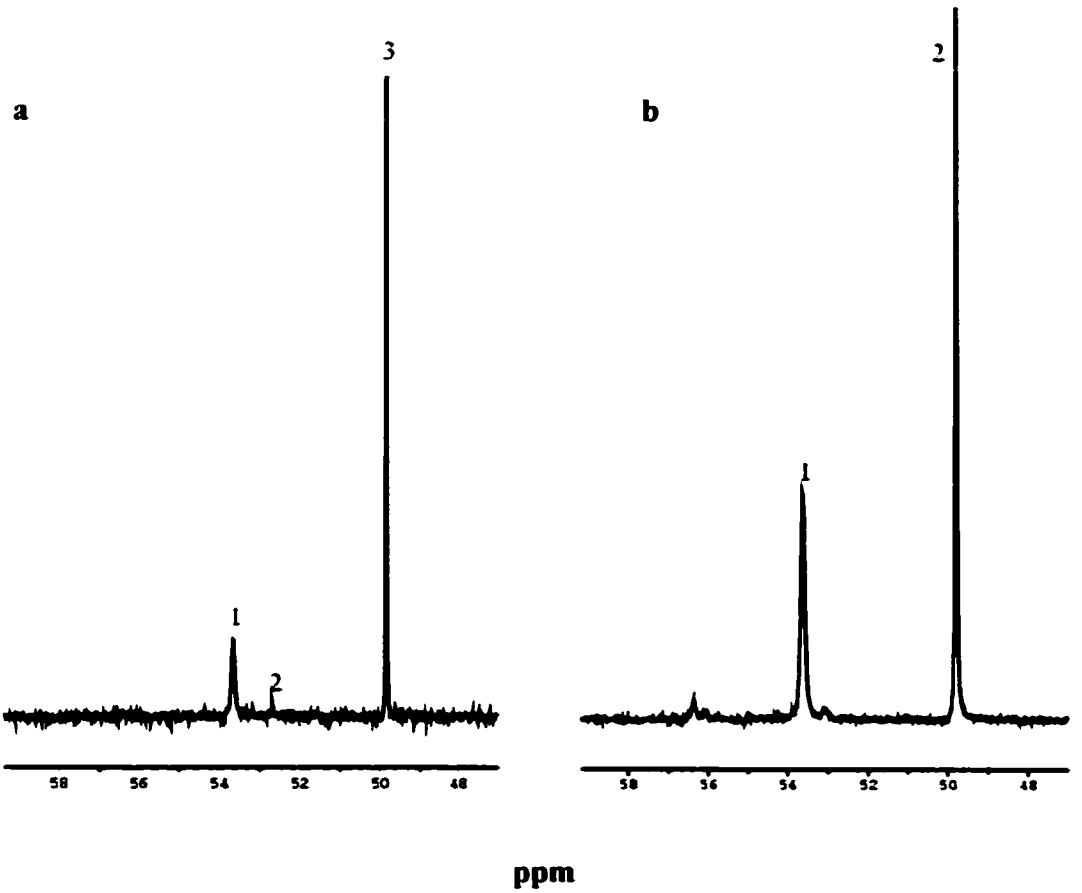


Figure 2-4. Expanded ^{13}C-NMR spectra of ^{13}C-methylated (LpH 6; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h) (a) ovalbumin and (b) cytochrome *c* after treatments with hydroxylamine, base and heat; (c) human albumin before hydroxylamine treatment and (d) after *in vacuo* hydroxylamine treatment for 4 h.

For ovalbumin (a), peak resonances correspond to the ^{13}C-enriched methyl groups in (1) $\mathbf{Me_3N^{\varepsilon+}}$-Lys (53.66 ppm); (2) unidentified α-$^{+}$N**Me$_3$** impurity (52.68 ppm); and (3) **MeOH** (49.83 ppm). The spectrum for untreated (with hydroxylamine) methylated ovalbumin contained resonances for α-, γ-, and δ-carboxyl ^{13}C-methyl esters at 53.76, 53.49, and 53.20 ppm, respectively.

For cytochrome *c* (b), peak resonances correspond to the ^{13}C-enriched methyl groups in (1) $\mathbf{Me_3N^{\varepsilon+}}$-Lys (53.61 ppm); and (2) **MeOH** (49.77 ppm). The spectrum for untreated (without hydroxylamine) methylated cytochrome *c* contained resonances for α-, γ-, and δ-carboxyl ^{13}C-methyl esters at 53.82, (53.49, 53.45), and (53.19, 53.15) ppm, respectively.

Continued...

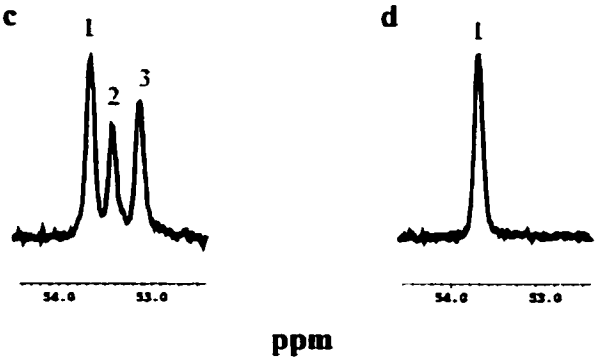


Figure 2-4. For human albumin (c) and (d), peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) $\mathbf{Me_3}N^{\varepsilon+}$-Lys (53.71 ppm); (2) C_αOO**Me** (53.66 ppm); (3) C_γOO**Me** (53.42 ppm); (4) C_δOO**Me** (53.13 ppm).

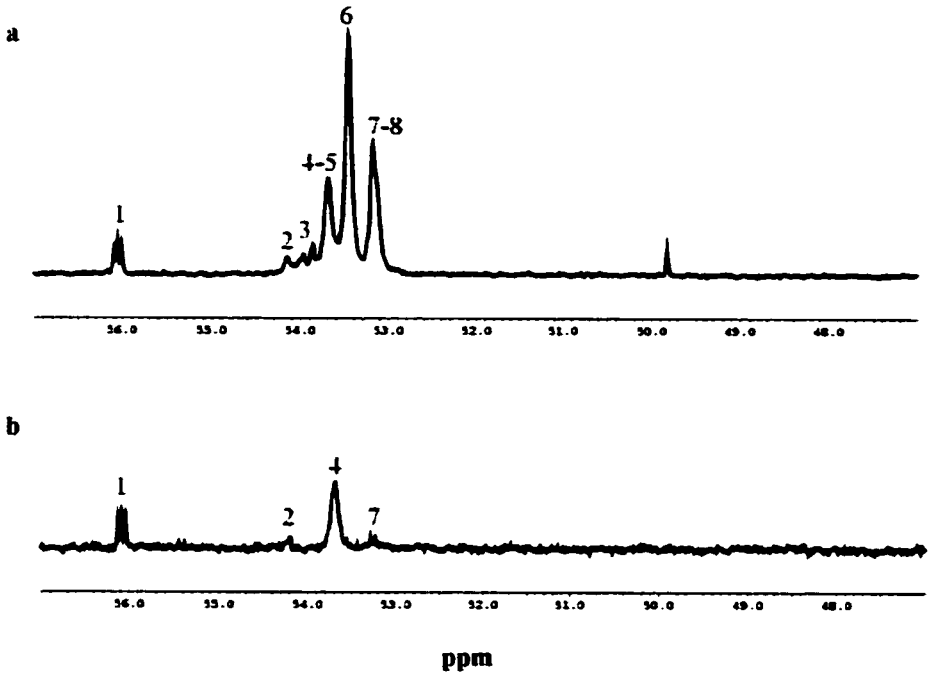


Figure 2-5. Expanded ^{13}C-NMR spectra of ^{13}C-methylated (LpH 7; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h) (a) soybean trypsin inhibitor before hydroxylamine treatment and (b) after *in vacuo* hydroxylamine treatment for 4 h.

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) tetramethylammonium cation (56.06 ppm triplet); (2) unidentified α-$^{+}$N**Me$_3$**, impurity (54.19 ppm); (3) C_αOO**Me** minor impurities (53.8-54.0 ppm); (4) **Me$_3$**N$^{\varepsilon+}$-Lys (53.68 ppm); (5) C_αOO**Me** (53.66 ppm); (6) C_γOO**Me** (53.46 ppm); (7) Asp-N$^{\alpha+}$**Me$_3$** (53.28 ppm); (8) C_δOO**Me** (53.18 ppm).

3.2.2 ^{13}C-Methylated Protein Hydrolysates

In the case of proteins with a blocked N-terminus, such as cytochrome c, and ovalbumin, the ^{13}C-NMR spectra of their hydrolysates should yield a single resonance corresponding to the resonance of the ^{13}C-trimethyl ε-amino group of N^{ε}-trimethyl Lys. Figure 2-6 (a, b) shows the main peak as a triplet at 53.65 ppm that corresponded to a ^{13}C-trimethlyl ε-amino Lys standard. Minor resonances were observed, which resulted from the derivatization of the N-terminal residues of minor protein impurities and possibly from incomplete hydrolysis. However, in proteins with a free N-terminal Asp residue, such as in the case of human albumin (Figure 2-6c), it is difficult to detect a resonance for the ^{13}C-trimethyl α-amino group and only that for the ^{13}C-trimethyl ε-amino groups of Lys residues (triplet) may be observed.

Aprotinin has one polypeptide chain with an N-terminal Arg residue (Dayhoff, 1972). The ^{13}C NMR spectrum of its hydrolysate is shown in Figure 2-6d. The resonance for the ^{13}C-trimethyl α-amino Arg derivative was observed at 52.59 ppm as a triplet in addition to that for the triplet corresponding to the ^{13}C-trimethyl ε-amino groups of Lys residues.

Chymotrypsinogen A has one polypeptide chain with an N-terminal half-cystine (Boyer, 1971) and its methylated hydrolysate was performic acid oxidized to yield the ^{13}C-trimethyl α-amino cysteic acid derivative. The latter amino acid derivative was observed in the ^{13}C-NMR spectrum as a small broad peak at 52.99 ppm in addition to the triplet for the ^{13}C-trimethyl ε-amino groups of N^{ε}-trimethyl Lys (Figure 2-6e).

The ^{13}C-NMR spectra of the 24 h acid hydrolysates of insulin ^{13}C-methylated at LpH 7.5 and 10 are shown in Figure 2-7a. The resonances derived from the trimethylated

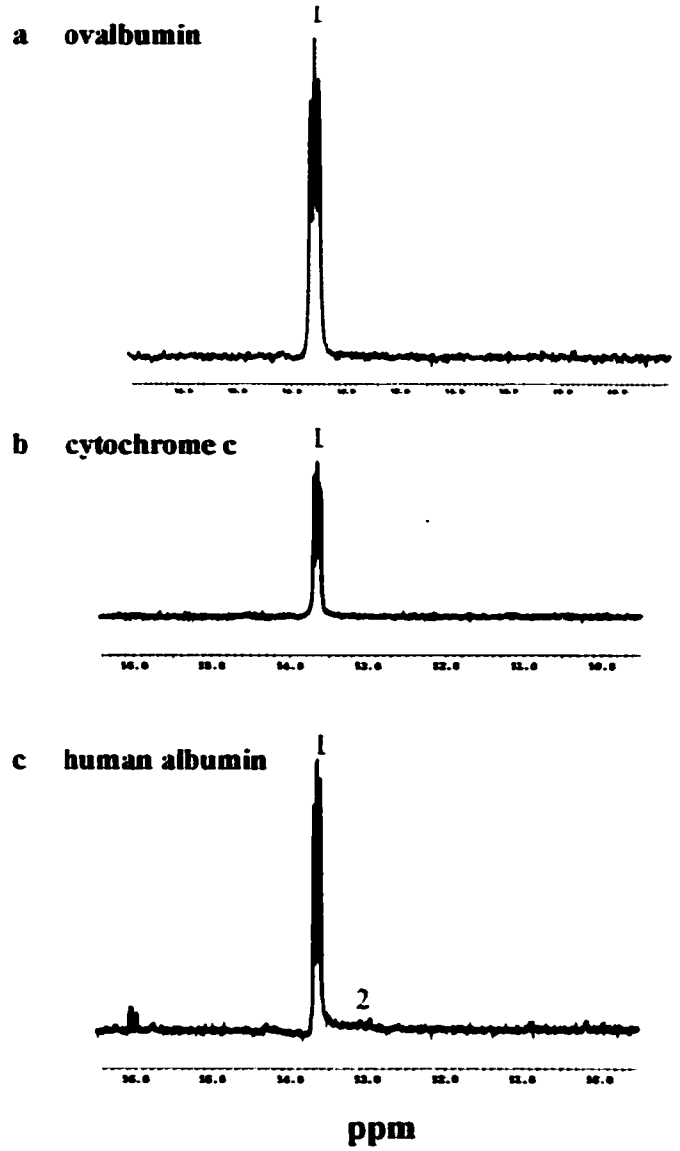


Figure 2-6. Expanded ^{13}C-NMR spectra (NT = 1024 scans) of ^{13}C-methylated protein hydrolysates (LpH 7; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h, followed by hydrolysis *in vacuo* with 6 N HCl) (a) ovalbumin, (b) cytochrome c, and (c) human albumin.

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

For ovalbumin (a), (1) $\mathbf{Me_3N^{\varepsilon+}}$-Lys.

For cytochrome c (b), (1) $\mathbf{Me_3N^{\varepsilon+}}$-Lys.

For human albumin (c), (1) $\mathbf{Me_3N^{\varepsilon+}}$-Lys, (2) $\mathbf{Me_3N^{\alpha+}}$-Asp.

Continued...

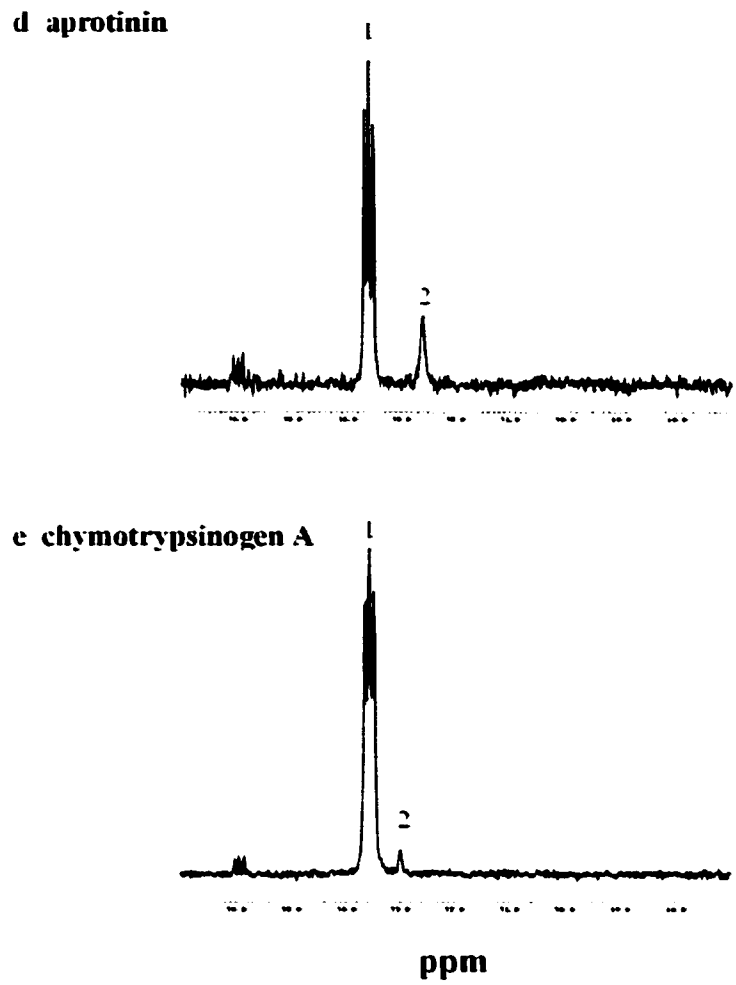


Figure 2-6. Expanded ^{13}C-NMR spectra (NT = 1024 scans) of ^{13}C-methylated protein hydrolysates (LpH 7; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h, followed by hydrolysis *in vacuo* with 6 N HCl) (d) aprotinin and (e) chymotrypsinogen A (performic acid oxidized).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

For aprotonin (d). (1) $\mathbf{Me_3N^{\varepsilon+}}$-Lys. (2) $\mathbf{Me_3N^{\alpha+}}$-Arg.

For chymotrypsinogen A (e). (1) $\mathbf{Me_3N^{\varepsilon+}}$-Lys. (2) $\mathbf{Me_3N^{\alpha+}}$-Cya (cysteic acid).

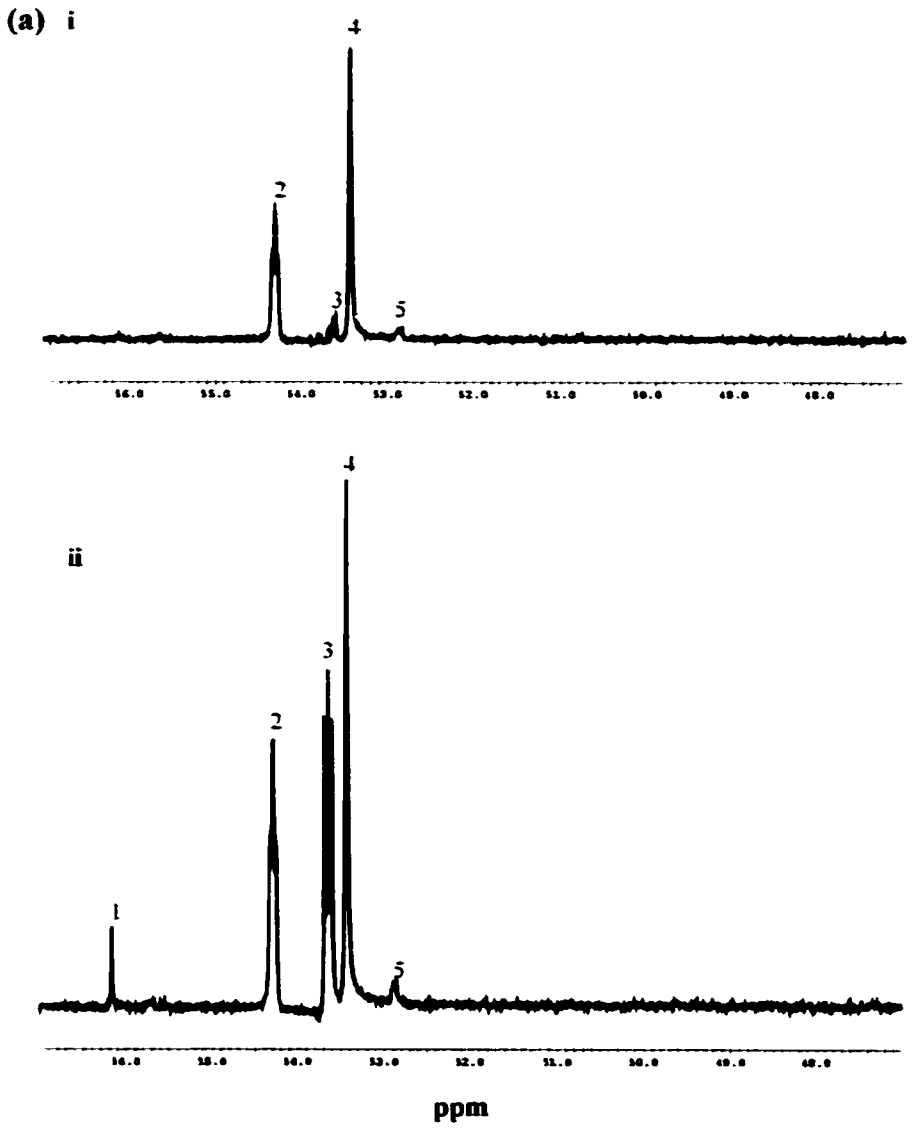


Figure 2-7(a). Expanded ^{13}C-NMR spectra (NT = 1024 scans) of ^{13}C-methylated insulin hydrolysates (a) at LpH 7.5 (i) and LpH 10 (ii) (reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h, followed by hydrolysis *in vacuo* with 6 N HCl).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(a), (1) unidentified; (2) $\mathbf{Me_3N^{\alpha+}}$-Gly; (3) $\mathbf{Me_3N^{\varepsilon+}}$-Lys; (4) $\mathbf{Me_3N^{\alpha+}}$-Phe-Val;

(5) $\mathbf{Me_3N^{\alpha+}}$-Phe.

Continued...

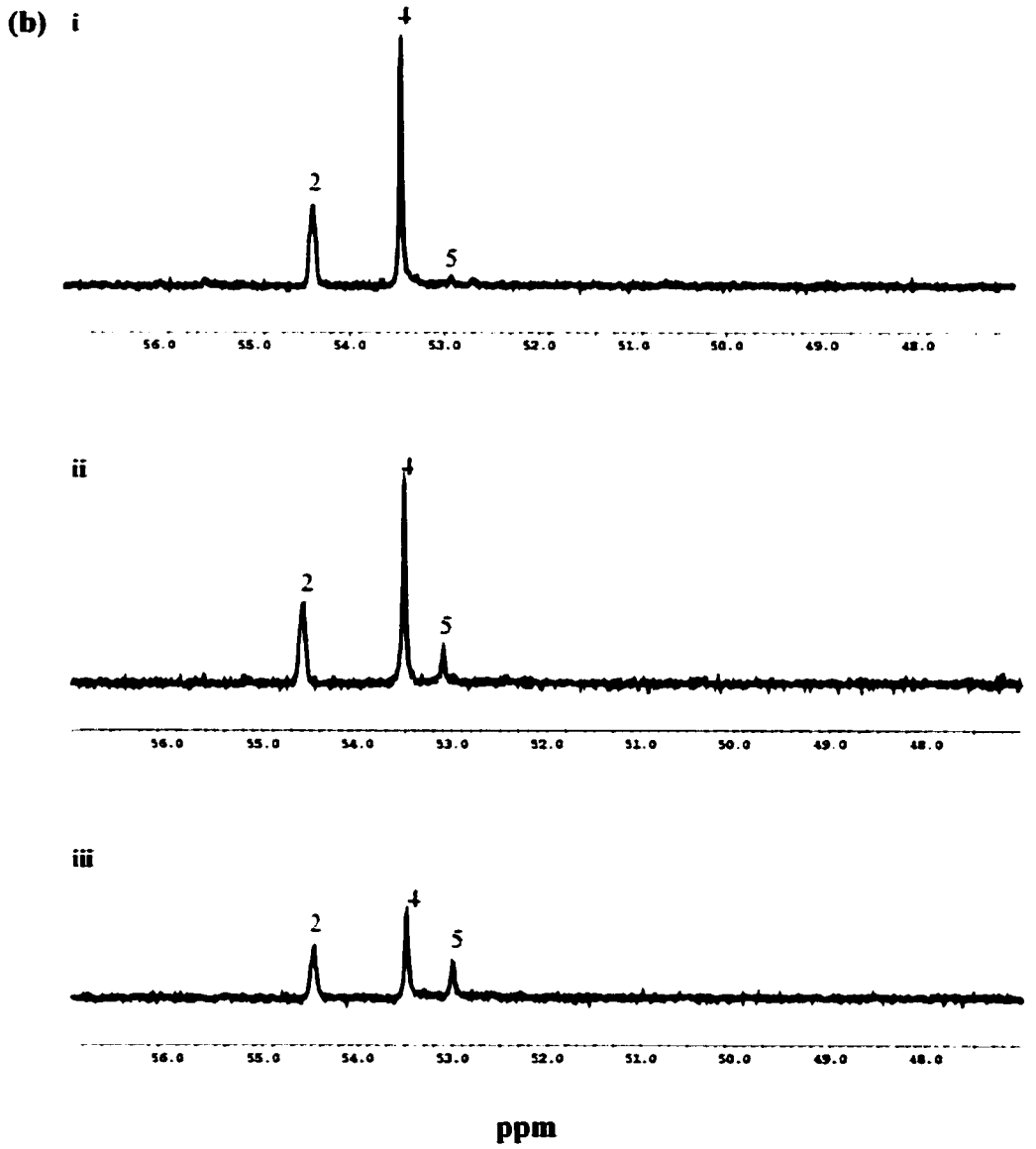


Figure 2-7(b). Expanded ^{13}C-NMR spectra (NT = 1024 scans) of ^{13}C-methylated insulin hydrolysates at LpH 7, observing (b) the effect of acid hydrolysis time (6 N HCl) on the extent of hydrolysis of Me_3-$N^{\alpha+}$-Phe-Val amide bond, (i) 24 h, (ii) 24+72 h, (iii) 24 h + 72 h + 8 d.

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(b), (1) unidentified; (2) $Me_3N^{\alpha+}$-Gly; (3) $Me_3N^{\varepsilon+}$-Lys; (4) $Me_3N^{\alpha+}$-Phe-Val;

(5) $Me_3N^{\alpha+}$-Phe.

α-amino-terminal Gly and Phe residues and the trimethylated ε-amino group of the single Lys residue are present at LpH 10, but the main resonances present at LpH 7.5 are those of the trimethylated α-amino-termini and the triplet derived from the trimethyl ε-amino group of Lys residue is present in a minor amount. The resonance at 53.5 ppm is derived from N^{α}-trimethylamino-Phe-Val, which is only slowly hydrolyzed to N^{α}-trimethyl-amino-Phe whose resonance is observed at 53.1 ppm as a triplet. At longer hydrolysis times (3 and 8 days; see Figure 2-7b) the intensity of the resonance at 53.1 ppm increases while that at 53.5 ppm decreases. ^{13}C-Trimethylated PheGlyGly tripeptide standard was hydrolyzed for 1, 3 and 8 days and yielded supporting evidence for the incomplete hydrolysis result observed for the methylated insulin hydrolysate.

Figure 2-8 shows the ^{13}C-NMR spectrum of α-chymotrypsin after ^{13}C-methylation at LpH 7.5 and 10, acid hydrolysis and performic acid oxidation. The multiple resonances in the amino region show that the chymotrypsin sample has undergone autolysis and that acid hydrolysis was likely incomplete. It was possible to assign resonances for trimethyl ε-amino Lys, trimethyl α-amino Ala, trimethyl α-amino Ile, trimethyl α-amino Ile-Val dipeptide, and trimethyl α-amino cysteic acid.

Figure 2-9 shows the ^{13}C-NMR spectrum of *Bacillus thuringiensis* toxin after ^{13}C-methylation at LpH 10 and acid hydrolysis. A small resonance at 52.8 ppm is due to the ^{13}C-trimethyl α-amino group of the N-terminal Ile amino acid (Bietlot *et al.*, 1989). Besides the resonances for the trimethyl ε-amino groups derived from lysyl residues at 53.7 ppm in the hydrolysate, there are distinct impurities in the protein preparation as indicated by the extra resonances present in the spectrum of the hydrolysate sample. It is possible that the sample was incompletely hydrolyzed.

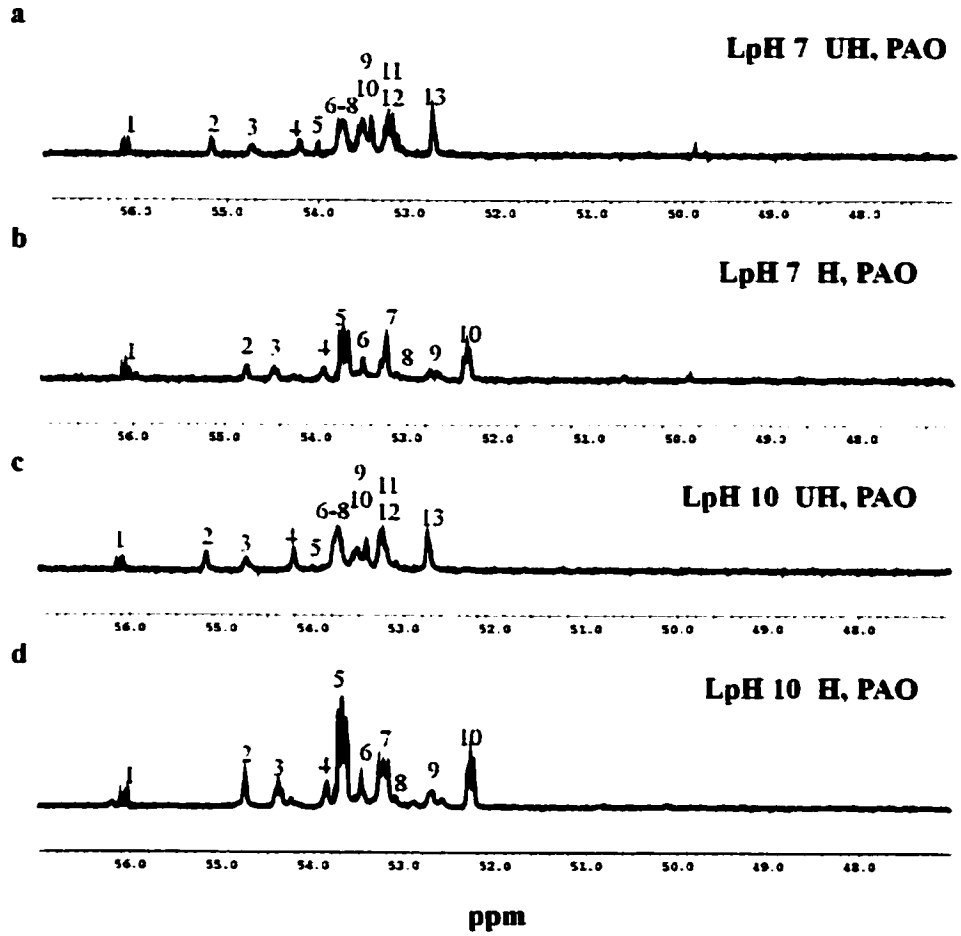


Figure 2-8. Expanded ^{13}C-NMR spectra (NT = 1024 scans) of ^{13}C-methylated α-chymotrypsin performic acid oxidized hydrolysates (reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h, followed by hydrolysis *in vacuo* with 6 N HCl) at LpH 7.5 (b) and LpH 10 (d) cf. unhydrolyzed samples (a), (c).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:
For (a) and (c), (1) Me_4-$^{+}$N Cl^- or I^- (56.1 ppm); (2), (3), (4) unidentified α-$^{+}$N**Me$_3$**, impurities (55.16 ppm, 54.69 ppm, and 54.21 ppm); (5) C_αOO**Me** (53.95 ppm; C-terminal Asn-245); (6) C_αOO**Me** (53.77 ppm); (7) **Me$_3$N$^{\varepsilon+}$**-Lys (53.68 ppm); (8) C_αOO**Me** (53.72 ppm); (6, 8) C-termini: Leu-13, Tyr-146; (9) C_γOO**Me** (53.50 ppm); (10) cysteic acid, Cya-N$^{\alpha+}$**Me$_3$** (53.45 ppm); (11) C_δOO**Me** (53.21 ppm); (12) **Me$_3$N$^{\alpha+}$**-Ile (53.23 ppm); (13) **Me$_3$N$^{\alpha+}$**-Ala (52.71 ppm). For (b) and (d), (1) Me_4-$^{+}$N Cl^- or I^- (56.1 ppm); (2), (3), (4) unidentified α-$^{+}$N**Me$_3$**, impurities; (5) **Me$_3$N$^{\varepsilon+}$**-Lys; (6) unidentified α-$^{+}$N**Me$_3$**, impurity; (7) **Me$_3$N$^{\alpha+}$**-Ile-Val; (8) cysteic acid, Cya-N$^{\alpha+}$**Me$_3$**; (9) **Me$_3$N$^{\alpha+}$**-Ile; (10) **Me$_3$N$^{\alpha+}$**-Ala.

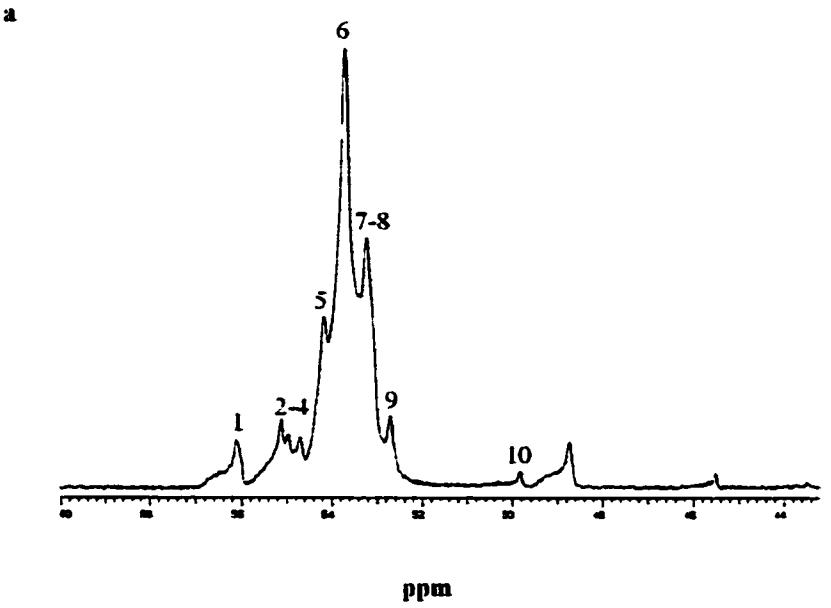


Figure 2-9(a). Expanded ^{13}C-NMR spectra of ^{13}C-methylated insecticidal toxin from *Bacillus thuringiensis* (a) (14208 scans) (reacted *in vacuo* with ^{13}C-iodomethane at LpH10).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

For (a), (1) Tyr-O**Me**; (2), (3), (4) **Me**$_3$N$^{\alpha+}$-impurities; (5) C$_\alpha$OO**Me**; **Me**$_3$N$^{\alpha+}$-impurity; (6) **Me**$_3$N$^{\varepsilon+}$Lys; (7), (8) C$_{\gamma,\delta}$OO**Me**; (9) **Me**$_3$N$^{\alpha+}$Ile; (10) MeOH.

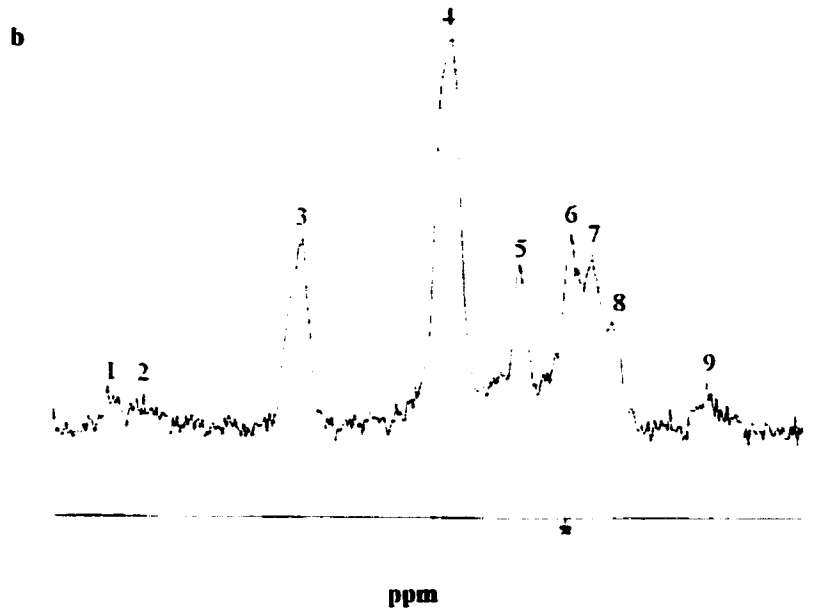


Figure 2-9(b). Expanded ^{13}C-NMR spectra of ^{13}C-methylated insecticidal toxin from *Bacillus thuringiensis* hydrolysate (b) (28438 scans) (LpH10).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

For (b), (1)-(3), (5)-(8) $Me_3N^{\alpha+}$-impurities or peptides originating from the N-terminus due to incomplete hydrolysis [e.g., $Me_3N^{\alpha+}$IleGlu, (6)-(8)]; (4) $Me_3N^{\varepsilon+}$Lys; (9) $Me_3N^{\alpha+}$Ile.

4. Discussion

The pH memory effect observed for the activity of lyophilized enzymes in nonaqueous solvents (Zaks and Klibanov, 1988) was interpreted as demonstrating that ionizable residues in the active site retained the ionization state they had in the solution from which they were prepared. Implicit in this explanation is that the pH memory effect applies to all ionizable groups in proteins. The observed ionization constants for α- and ε-amino groups in proteins are 6.8 - 8.0 and 10.4 - 11. 1, respectively (Creighton, 1993a). For proteins lyophilized at pH values between 7 and 8.5, it is therefore expected that the ε-amino groups will be completely protonated and nonreactive whereas α-amino groups will be substantially deprotonated and reactive. Under nonaqueous conditions, the α-amino groups should react preferentially with electrophilic reagents such as iodomethane. The results obtained in the present study for the methylation of amino groups indeed show that advantage can be taken of the pH memory effect in lyophilized proteins to achieve a selectivity that is unattainable under aqueous conditions. By judicious selection of the pH of lyophilization of a protein, its α-amino group can be modified while the ε-amino groups are modified to only a minor extent in some proteins or not at all in others.

During the initial stages of this study, a sharp resonance due to ^{13}C-methanol at 49.8 ppm was observed in some spectra. Further investigation showed that the size of this peak was directly related to the time the sample remained in solution prior to ^{13}C-NMR analysis. This observation suggested that methyl esters of carboxyl groups were formed by reaction with iodomethane under the nonaqueous conditions. It is important to note that samples were analyzed at pD 7.9 prior to the discovery that

carboxyl methyl esters are products of the *in vacuo* methylation reaction. On standing in the aqueous solution at neutral or slightly alkaline pH values prior to NMR analysis, these esters were being hydrolyzed giving rise to the ^{13}C-methanol resonance. Incubation of the methylated protein with hydroxylamine, also liberated ^{13}C-methanol concomitant with the disappearance of several of the observed resonances, e.g., Figures 2-1b and 2-1d, confirming the presence of methyl esters. The presence of these ^{13}C-resonances from methylation of carboxyl groups is a complicating factor as they occur in the same spectral region as ^{13}C-trimethylamino groups. However, they are easily removed from the spectrum by alkaline treatment or addition of hydroxylamine along with mild heating to accelerate the reaction, leaving only the resonances from the trimethylated amino groups of interest in the region from 52 to 57 ppm.

In some cases, the removal of the methyl ester resonances from the ^{13}C-NMR spectrum was not complete with either hydroxylamine or base treatment or both in aqueous medium. The gelling of some proteins (α-chymotrypsin, ovalbumin and human albumin) after hydroxylamine treatment resulted in the incomplete conversion of methyl esters to hydroxamates. The complete removal of carboxyl methyl ester resonances from the ^{13}C-NMR spectra of methylated proteins via the *in vacuo* hydroxylamine reaction, as was carried out on ^{13}C-methylated ribonuclease A, α-chymotrypsin and soybean trypsin inhibitor (see Chapter 2), proved invaluable for our investigation (see Sections 2.5.1, 2.5.2, and 3.1).

The positive results of the *in vacuo* hydroxylamine reaction led us in the direction of improving upon the colorimetric technique for the quantification of methyl esters. From the results observed in this study in terms of the difficulty of methyl ester removal,

the method of Halpin and Richardson (1985) is clearly unsatisfactory because their protocol yields incomplete conversion of methyl esters to hydroxamates via the selective reaction of hydroxylamine at the sites of esterification. The latter reaction yields hydroxamic acid with the release of methanol. In an acidic environment, excess ferric ion will form a 1:1 complex with hydroxamic acid leading to a reddish-purple chelate that can be quantified spectrophotometrically. From the stoichiometry of the various reactions, the absorbance of the coloured chelate should be directly related to the number of esterified sites. Thus, with the Halpin and Richardson (1985) protocol, the colorimetric reaction would give an erroneous result. The results of the present study with the *in vacuo* hydroxylamine reaction can lead to an improved accuracy in the colorimetric method for the determination of the extent of esterification of carboxyl groups, be it in protein or lipid molecules (Halpin and Richardson, 1985; Kates, 1986).

While the resonances of individual ^{13}C-trimethylated α-amino groups occur over about a 5 ppm range, those from individual ^{13}C-trimethylated ε-amino groups appear as a single peak. However, the vast majority of proteins contain very few α-amino groups, usually one or two, and without preferential modification of α-amino groups the resonance from the ^{13}C-trimethylated ε-amino groups would dwarf that from a solitary ^{13}C-trimethylated α-amino group. This would make it more difficult to observe the individual resonances from ^{13}C-trimethylated α-amino groups, especially when they arise from proteins that are present as minor impurities or when they lie close to that of the ^{13}C-trimethyl ε-amino group(s).

In addition to amino groups, several other groups in the protein react with ^{13}C-iodomethane at LpH values between 6 and 7. As already noted, carboxyl groups

react to form methyl esters, but this does not present a problem because they are easily hydrolyzed, and thus, analysis of the ^{13}C-NMR spectrum can be simplified. Imidazole side-chains of histidine residues form the dimethylated imidazolium derivative and the methionine side-chain forms the dimethylsulfonium derivative (Taralp and Kaplan, 1997). The ^{13}C-resonances of these derivatives lay far upfield from those of the trimethylamino groups, i.e. 25 ppm to 37 ppm, and do not interfere with the analysis of the trimethylated amino groups.

It was possible that the conditions of the *in vacuo* methylation reaction at high LpH values would give rise to nucleophilic substitution or Hofmann elimination reactions (March, 1985). A nucleophilic substitution reaction can take place if the quaternary ammonium has a OH^- counterion; i.e., at high pD) to give an alcohol ($R_3R'N^+OH^- \rightarrow ROH + (CH_3)_2N\text{-}R'$; e.g., dimethylamino Phe or Gly N-termini of methylated insulin; should observe a new peak in the 40-50 ppm range). When none of the four alkyl groups on the nitrogen has a β-hydrogen, as is the case with trimethyl-α-amino-Gly-Protein in methylated insulin, then substitution is the only reaction possible. Hofmann elimination usually occurs via an E2 mechanism and Hofmann's rule, whereby the double bond goes mainly toward the least highly substituted carbon, is generally obeyed (March, 1985). For highly hindered molecules (e.g., trimethyl-α-amino-Phe-Protein in methylated insulin), a concerted five-membered Ei mechanism has been shown to operate (i.e., the OH^- does not attract the β-hydrogen, but instead removes one of the methyl hydrogens), which yields R_3N (i.e., trimethylamine (47.55 ppm) from trimethyl-α-amino-Phe-Protein in insulin) and the olefin product (e.g., PhCH=CH-CO-Protein, where Ph = phenyl group of Phe side-chain of the N-terminal amino acid residue in methylated insulin). It is

important to note that during the course of this study there was no evidence that indicated that either nucleophilic substitution or Hofmann elimination occurred under the reaction conditions employed. It was observed, however, that Hoffman elimination might have occurred with an insulin sample that was made accidentally highly alkaline (from the desired pD 5.4) during the preparation of the sample for NMR analysis, as a peak at 48.7 ppm was observed. Whether trimethylamine was formed and the expected resonance of 47.5 ppm (Breitmaier and Voelter, 1987) for the elimination product occurred closer to the observed 48.7 ppm under the conditions that the sample was prepared still remains to be determined; however, it is safe to conclude that either nucleophilic substitution or Hofmann elimination does not occur during the *in vacuo* methylation reaction.

It should be noted that at LpH values between 6 and 7 the α-amino groups are not completely methylated. If an amino group has a pK_a of 8 then at LpH 7, it is expected, on the basis of the pH memory effect in lyophilized proteins that approximately 10% of the amino groups would be methylated. Methylation at higher LpH values could be used to increase the sensitivity of the approach, but this would increase the degree of trimethylation of the ε-amino groups and could, as noted above, obscure the resonances of trimethylated α-amino groups. However, in specific cases, it may be advantageous to achieve a greater sensitivity by carrying out the methylation reaction at higher LpH values. In general, our experience indicates that an LpH value of 7 would be optimal for most proteins, as there is significant methylation of the α-amino groups with minimal methylation of the ε-amino groups.

Multiple resonances from ^{13}C-trimethylamino groups can arise for several reasons. The protein itself could consist of several polypeptide chains with different amino-termini. In such a case, the resonances should be of comparable size and further verification could be obtained by methylating the protein at a higher LpH value where the differences in the extent of methylation of α-amino groups due to differences in their pK_a values will be minimized. Proteolysis or minor impurities could also account for multiple resonances. However, in such cases it is expected that the intensities of these peaks will be much smaller than those from the major components. While such impurities can be detected by SDS-PAGE, they may be missed if, as is often the case, the impurities and the major component are of similar size.

It is expected that for proteins with a blocked amino-terminus no resonance from an α-amino group will be detected. This is what was observed with the commercial samples of proteins with blocked amino-termini used in this investigation (Figures 2-4 and 2-6). At present, there is no method for determining the presence of a blocked amino-terminus other than the failure to detect an amino-terminus using conventional N-terminal analysis or the failure to obtain a sequence on an automated sequencer. Therefore, the approach described here provides an additional method for confirming the presence of a blocked amino-terminus in a protein.

Trimethylated amino groups are acid stable and, in principle, it should be possible to identify a trimethylated amino acid after acid hydrolysis and NMR spectroscopic analysis in much the same manner as conventional N-terminal analysis using fluorescent or coloured derivatives. However, in exploring this possibility, it was found that in many cases the hydrolysis product was the N-terminal ^{13}C-trimethylated dipeptide and not the

N-terminal ^{13}C-trimethylated amino acid. The trimethyl group evidently gives rise to a great deal of steric hindrance, and it was found that in some cases acid hydrolysis in 6 N HCl at 110°C for up to one week was required to obtain greater than 50% conversion of the ^{13}C-trimethylated dipeptide to the ^{13}C-trimethylated terminal amino acid for NMR spectroscopic analysis. While in principle, this approach could be used to determine the identity of the ^{13}C-trimethylated amino acid from the unique position of its resonance, the long hydrolysis time required makes this approach impractical. It should be noted however, that the position of the resonances of the ^{13}C-trimethylated dipeptides are relatively close, i.e. within 0.5 ppm (downfield), to that of the free ^{13}C-trimethylated terminal amino acid (Table 2-1), and both are also relatively close to the resonance observed in the intact protein. For example, the resonance of the ^{13}C-trimethylated phenylalanine amino-terminus in intact insulin is observed at 53.4 ppm, that for the ^{13}C-trimethylated Phe-Val at 53.5 ppm, and ^{13}C-trimethylated Phe at 53.1 ppm. While it is not possible in general to make an unequivocal identification of the N-terminal amino acid on this basis, it is relatively easy to determine if the resonance is consistent with a specific N-terminal amino acid. In the case of proteins where the N-terminal amino acid is known this can be useful confirmatory evidence. In other cases, it can be used to rule out a specific amino acid. For example, in the case of the minor resonance observed in ovalbumin (Figure 2-4a), it is possible to rule out that this resonance is due to a small amount of the unblocked protein because the resonance at 52.7 ppm is not consistent with a ^{13}C-trimethylated Gly-Leu which occurs at 55.1 ppm. It was on this basis that it was possible to assign the resonance observed at 55.1 ppm in methylated insulin (Figure 2-1a)

Table 2-1. Chemical Shifts of N^{α}-^{13}C-Trimethylated Derivatives of Amino Acids and Peptides

Amino Acid	Peptide or Amide	^{13}C-Resonance (ppm)*
Me_3N^+-Ala		52.2
	Me_3N^+-Ala-Ala	52.6
Me_3N^+-Arg		52.6
Me_3N^+-Asn		53.6
Me_3N^+-Asp		53.3
Me_3N^+-Cys-(SO_3^-)		53.1
Me_3N^+-Gln		52.7
Me_3N^+-Glu		52.6
Me_3N^+-Gly		54.3
	Me_3N^+-Gly-NH_2‡	54.9
	Me_3N^+-Gly-Leu†	55.1
Me_3N^+-His		53.0
Me_3N^+-Ile		52.5
	Me_3N^+-Ile-NH_2†	53.0
	Me_3N^+-Ile-Val-NH_2‡	53.1
Me_3N^+-Leu		52.5
Me_3N^+-Lys		52.6
Me_3N^+Met($^+S^{13}MeMe$)		52.6
Me_3N^+-Phe		53.1
	Me_3N^+-Phe-Gly-Gly†	53.4
Me_2N^+-Pro		53.0
Me_3N^+-Ser		53.8
Me_3N^+-Thr		54.1
Me_3N^+-Trp		53.1
Me_3N^+-Tyr		52.8
Me_3N^+-Val		52.6

* Chemical shift referenced to internal acetonitrile, at 1.70 ppm.
† Values taken from Taralp and Kaplan, 1997
‡ Values taken from Taralp, 1997

to the ^{13}C-trimethylated glycine amino-terminus and that at 53.4 ppm to the ^{13}C-trimethylated phenylalanine amino-terminus.

Although the results with the protein hydrolysates was satisfactory, it may be worthwhile to study the effect of other organic acids (e.g., trifluoroacetic acid) and higher temperatures (Creighton, 1993b; Hill, 1965) on the extent of peptide bond hydrolysis to verify conclusively that the acid hydrolysis approach presented herein is indeed an impractical one due to incomplete hydrolysis as described above.

The results with insulin clearly show the selectivity of the methylation reaction toward α- vs ε-amino groups for the protein hydrolysates in the same manner as was observed for the methylated intact proteins.

In the ^{13}C-NMR spectra of the protein hydrolysates, the coupling of the spin ½ carbon-13 atoms of the methyl groups in the trimethylamino derivatives of the amino acids to the quadrupolar nitrogen nucleus of spin 1 (Mann, 1978) yielded 1:1:1 triplets with coupling constants on the order of 12 Hz. Splitting was not observed for some derivatives due to the asymmetric environment surrounding the quadrupolar nitrogen atom, which leads to efficient quadrupolar relaxation (see Chapter 4, Section 1.2) of the nitrogen atom making the nitrogen atom unable to couple to methyl carbons in a manner that is observed on the NMR time scale. In the special case of extreme tetrahedral symmetry surrounding the nitrogen atom, efficient quadrupolar relaxation is not possible (Brevard and Granger, 1981). In the trimethylamino N-terminal Phe-Val dipeptide and the trimethylamino Phe-Gly-Gly tripeptide standard, the overall symmetry surrounding the nitrogen atom is reduced, and the splitting pattern is lost due to more efficient relaxation. Although Taralp (1997, Doctoral Thesis) did not identify the resonance for

the trimethylamino N-terminal Phe-Val, the observation that the resonance was not a triplet led to the suggestion that the resonance was not derived from a free trimethylamino acid derivative. Moreover, the change in its chemical shift value following hydrolysis (compared to the intact methylated insulin) was small, and Taralp did suggest that the resonance might have been derived from a trimethylamino dipeptide. However, Taralp suggested that the origin of the resonance at 53.5 ppm in methylated insulin may have arisen from an error in post-translational processing in the conversion of proinsulin to insulin. Despite this respectable reasoning, Taralp was unable to characterize the peak at 53.5 ppm in ^{13}C-methylated insulin hydrolysate.

As NMR spectrometers are widely available, the approach described herein provides a facile and relatively rapid method of obtaining evidence for the presence or absence of α-amino groups, the number of amino-termini and minor protein impurities in a protein preparation. High quality spectra are obtainable in less than 30 minutes on milligram quantities of protein using a 9.4 Tesla spectrometer (^{1}H, 400.13 MHz; ^{13}C, 100.6 MHz) with a DEPT-135 pulse sequence. If required, greater sensitivity could be achieved with longer acquisition times or with higher field instruments.

In summary, it has been demonstrated that nonaqueous methylation of proteins with ^{13}C-iodomethane can be used to trimethylate α-amino groups preferentially. The ^{13}C-trimethylated amino groups can be used as an NMR probe for determining the number of different α-amino groups in a protein preparation and for making a tentative identification of the amino-terminal amino acid or acids. The absence of a resonance from a ^{13}C-trimethylated α-amino group can be used to confirm the presence of a blocked amino-terminus. In addition, the presence of minor resonances from trimethylated

α-amino groups can provide evidence of protein impurities or proteolysis in protein preparations.

5. Acknowledgement

This research was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

6. References

Bailey, L. (1964) *Techniques in Protein Chemistry*, Elsevier Publishing Company, p. 99.

Bendall, M. R., Doddrell, D. M., Pegg, D. T., and Hull, W. E. (1982) In *DEPT, Bruker Analytische Messtechnik*, Karlsruhe.

Blundell, T., Dodson, G., Hodgkin, D. and Mercola, D. (1972) *Adv. Protein Chem.*, (Anfinsen, JR., C. B., Edsall, J. T., and Richards, F. M., eds.), pp. 371-375, and references therein, Academic Press, New York.

Boyer, P. D., ed. (1971) In *The Enzymes*, 3rd ed., Vol. **III**, pp. 165-249.

Breitmaier, E., and Voelter, W. (1987). In *Carbon-13 NMR Spectroscopy: High-Resolution Methods and Applications in Organic Chemistry and Biochemistry*, 3rd ed., VCH Publishers, New York. (a) p. 109; (b) p. 208; (c) p. 228; (d) p. 237.

Brevard, C. and Granger, P. (1981) *Handbook of High Resolution Multinuclear NMR*, John Wiley and Sons, New York, p. 104.

Creighton, T. E. (1993). *Proteins: Structure and Molecular Properties*, 2nd ed., W.H. Freeman and Company. (a) p. 6; (b) pp. 31-34

Darbre, A. (1986). *Practical Protein Chemistry* (Darbre, A., ed.), Wiley, New York. (a) p.81; (b) pp. 236-239

Dayhoff, M. O., ed. (1972) Atlas of Protein Sequence and Structure, Vol. 5, The National Biomedical Research Foundation, Silver Springs, Maryland.

Ferry, J. D. (1948) *Adv. Protein Chem.*, **4**: 1.

Glasoe, P. K. and Long, F. A. (1960) *J. Phys. Chem.*, **64**: 188-190.

Halpin, M. I. and Richardson, T. (1985). *J. Dairy Sci.*, **68**: 3189-3198.

Hames, B. D. (1990) In *Gel Electrophoresis of Proteins: A Practical Approach*, (Hames, B. D. and Rickwood, D., Eds.) IRL Press, Oxford, pp. 1-139.

Hill, R. L. (1965) *Adv. Protein Chem.*, **20**: 37-107.

Hirs, C. H. W. (1956) *J. Biol. Chem.*, **219**: 611.

Kates, M. (1986) *Laboratory Techniques in Biochemistry and Molecular Biology: Techniques of lipidology isolation, analysis and identification of lipids*, Burdon, R.H. and van Knippenberg, P.H., eds., Elsevier Science Publishers, 2nd edn., Vol. 3, Part 2, p.119-120.

Magdassi, S., ed. (1996) *Surface Activity of Proteins, Chemical and Physiochemical Modifications*, Marcel Dekker, Inc., p.28.

Mann, B.E. (1978) in *NMR and the Periodic Table*, (Harris, R.K. and Mann, B.E., eds.), Academic Press Inc. (London) Ltd. (a) p. 96; (b) p. 98.

March, J. (1985) In *Advanced Organic Chemistry*, Third Edition, Wiley-Interscience, New York, p.906-908.

Taralp, A. (1997) In *Aqueous and Nonaqueous Chemical Approaches to Elucidating Structure and Function in Proteins*. Ph.D. Thesis, Department of Chemistry, University of Ottawa, Ottawa, ON.

Taralp, A. and Kaplan, H. (1997). *J. Protein Chem.*, **16:** 183-193.

Zaks, A. and Klibanov, A. M. (1988). *J. Biol. Chem.*, **263:** 3194-3201.

Chapter 3

In Vacuo Esterification of Carboxyl Groups in Lyophilized Proteins

1. Introduction

In Chapter 2, the pH memory effect in lyophilized proteins (Zaks and Klibanov, 1988) was used to achieve preferential trimethylation of N-terminal α-amino groups. ^{13}C-Methylation of protein was carried out via the nonaqueous *in vacuo* technique of Taralp and Kaplan (1997). It was possible to use ^{13}C-NMR spectroscopy to identify whether a protein has a free or blocked amino-terminus, and in the case of a free amino-terminus, to make a tentative identification of the amino-terminal amino acid. In addition, the number and relative intensities of the amino-terminal resonances observed could be used to determine the homogeneity of the protein preparation. The success of identifying the N-terminal residue(s) of a protein as a preliminary criterion of protein identity and homogeneity was dependent on the identification and removal of the ^{13}C-resonances from the ^{13}C-NMR spectra that corresponded to the carboxyl ^{13}C-methyl ester derivatives (using the carboxyl methyl ester group-specific reagent, hydroxylamine (Jencks, 1958)). In the present chapter, a thorough investigation is conducted with several proteins to characterize the *in vacuo* methylated derivatives of carboxyl groups formed by reaction with ^{13}C-iodomethane. In principle, the same approach as that carried out for the identification of N-terminal residue(s) could be used to identify C-terminal α-carboxyl groups, but no cost-effective aqueous chemical modification procedure exists for the introduction of a ^{13}C-label into carboxyl groups. Unlike the chemical modifications carried out in aqueous media, the nonaqueous protocol

requires the use of only small amounts of reagent, so it is practical to use ^{13}C- or ^{14}C-labeled reagents, which would otherwise be prohibitively costly.

As mentioned in the introductory chapter, the studies of Taralp and Kaplan (1997) did not reveal that ^{13}C-iodomethane forms stable derivatives with the carboxyl function of aspartyl, glutamyl, or C-terminal residues. However, the authors did note the presence of a small resonance corresponding to ^{13}C-methanol at 49.79 ppm, but had attributed it to being due to traces of water reacting with iodomethane. It is not surprising that Taralp and Kaplan (1997) did not identify the ^{13}C-methyl esters under the conditions employed. Protein was reacted *in vacuo* at LpH 7.5 and LpH 10, where the resonances of the ^{13}C-trimethylammonium derivatives of N^{ϵ}-amino groups of lysyl residues overlap and may obscure the observation of the carboxyl ^{13}C-methyl ester resonances. Moreover, the methyl esters have a lower sensitivity of detection compared to the trimethylated amino groups since they are enriched with only one ^{13}C atom as opposed to the three that the latter derivatives possess.

However, technological advances in ^{13}C NMR spectroscopy have greatly enhanced its sensitivity. In particular, the availability of a higher magnetic field instrument than that used in the original work of Taralp and Kaplan (1997), and the employment of the DEPT-135 pulse sequence (Bendall *et al.*, 1982) have provided us the opportunity to further characterize the *in vacuo* reaction of ^{13}C-iodomethane with several proteins.

In the present study, it is demonstrated that the conditions of the *in vacuo* nonaqueous methylation procedure can be geared toward the preferential reactivity of protein carboxyl groups by simply adjusting the pH of lyophilization (LpH). Herein the

utility of iodomethane as a good protein-modifying reagent is further exemplified by the use of the pH memory effect of lyophilized proteins (Zaks and Klibanov, 1988), as it can effect different and selective or preferential chemical modifications with distinct functional groups under appropriate conditions. The current NMR technology provides sufficient sensitivity for the identification of the ^{13}C-chemical shifts (Breitmaier and Voelter, 1987c) of the different types of methyl esters that can be formed by reacting protein with methylating reagents (i.e., α-, γ-, and δ-carboxyl groups of C-terminal, aspartyl, and glutamyl amino acid residues in proteins, respectively). Furthermore, the extent of methylation of free carboxyl groups in protein using iodomethane was estimated by difference titration and a comparison is made with a novel *in vacuo* esterification reaction using ^{13}C-methanol and hydrogen chloride catalyst. To our knowledge, the latter modification, as carried out *in vacuo* on lyophilized protein under nonaqueous conditions, has not been performed previously.

The site-specific modification of carboxyl groups in proteins is somewhat difficult to achieve, as is differentiation between aspartyl and glutamyl residues (Lundblad, 1995a). A wide variety of reagents have been employed in the chemical modification of carboxyl groups in proteins under aqueous conditions with carbodiimide-mediated modification being the most widely used procedure (Hoare and Koshland, 1966; Means and Feeney, 1971; Lundblad, 1995a). Besides iodomethane (Blackburn, Hemphill Carter, and Phillips, 1941) and acidic methanol (Fraenkel-Conrat and Olcott, 1945; Broomfield *et al.*, 1965; Wilcox, 1967; Means and Feeney, 1971), other protein carboxyl group methylating reagents of varying specificity are: methyl sulfate (Blackburn, Hemphill Carter, and Phillips, 1941), dimethyl sulfate (Wissmann and Hillen, 1991),

diazomethane (Rutherford, Patterson, and Harris, 1940; Herriott, 1947), trimethyloxonium tetrafluoroborate (Paterson and Knowles, 1972), triethyloxonium tetrafluoroborate (Parsons *et al.*, 1969; Parsons and Raftery, 1969), *N*-ethyl-5-phenyl-isoxazolium-3'-sulfonate (Woodward's Reagent K; Woodward *et al.*, 1961; Woodward and Olofson, 1966) and other isoxazolium salts (Means and Feeney, 1971; Lundblad, 1995a). Other than the standard acid-catalyzed esterification by suspension of the protein in acidic alcohol (Means and Feeney, 1971), there has been no facile procedure reported for the esterification of carboxyl groups. *In vacuo* chemical modification of lyophilized proteins has many advantages (Taralp and Kaplan, 1997; see Chapter 1, Section 2.2). One of the objectives of the present investigation was to determine if esterification of carboxyl groups in proteins could be achieved efficiently and selectively by *in vacuo* chemical modification. Another objective of the present investigation was to determine whether ^{13}C-NMR spectroscopy could be used to study carboxyl group modifications in proteins and have practical applications in the characterization of proteins.

Free α-, γ- or δ-carboxyl groups derived from α-carboxyl terminal, β-aspartyl, and γ-glutamyl amino acid residues in protein molecules provide the anionic groups that participate in the formation of salt bridges with counterpart cation and/or hydrogen bonds (Perutz, 1978). Neutralization of the negative charges by carboxyl methyl-esterification is, therefore, expected to induce changes in the conformation of protein structure as well as in solvation properties of the protein. Thus, methyl-esterification of specific carboxyl group(s) could most likely cause specific functional modifications of the protein. Proteins of interest can be chemically methylated or esterified at varying extents and the

present *in vacuo* nonaqueous methodology can be exploited for investigating structure-function relationships.

In vivo, methyl-esterification of free carboxyl groups in cellular protein substrate is catalyzed by protein methylase II (protein carboxyl methylase; *S*-adenosylmethionine: protein carboxyl *O*-methyltransferase; EC 2.1.1.24), yielding protein methyl ester (Paik and Kim, 1980b). Esterification of protein carboxyl groups is believed to play a very crucial role in chemotactic and secretory processes, which are as yet not well-understood phenomena. Table 3-1 lists various experimental model systems for investigating the biological significance of the protein carboxyl methylation reaction. It is hypothesized that the enzymatic methyl esterification neutralizes vesicular membrane protein and consequently induces vesicular fusion, thus rendering excitation-secretion coupling of vesicular components (Paik and Kim, 1980c).

There has been difficulty in studying enzymatically methyl-esterified proteins because they may lose their modification during the purification procedure or during bioactivity assays due to the labile nature of the carboxyl methyl esters. Interestingly, it has been shown that enzymatically prepared protein methyl esters are 25 times more labile than those that are chemically synthesized (Paik and Kim, 1980a). It is clear that the overall structure of the protein plays an important role in determining the stability of an ester derived from it. It is highly conceivable that the neighbouring amino acid side-chains influence greatly the stability of the ester linkage, analogous to the situation where the stability of the amide bond in the protein molecule is influenced by the neighbouring amino acid side-chains (Robinson, 1974). In fact, Kim and Paik (1976) have shown that enzymatically synthesized protein methyl ester occurs at specific sites such that

ionization of neighbouring amino acid side-chains enhances the rate of hydrolysis. Moreover, Johnson and Aswad (1985) have shown that nonenzymatic cyclic imide formation immediately proceeds demethylation of transiently methylated adrenocorticotropin (by brain protein carboxyl methyltransferase) and can explain the unusual lability of mammalian protein methyl esters in general. The evidence includes: (1) the rapid rate of methyl ester hydrolysis, which is consistent with intramolecular catalysis, (2) the inability of the demethylated product to be remethylated, (3) the charge of the product, and (4) its rate of breakdown. These findings suggest that protein carboxyl methylation in mammalian tissues is not a simple on/off reversible modification. The authors suggested that carboxyl methylation may serve to activate selected protein carboxyl groups for subsequent longer lasting modifications, possibly subserving a role in protein repair, degradation, crosslinking, or some other as yet undiscovered alteration of protein structure.

1.1 Focus of the Present Research

In the present research it is demonstrated that iodomethane is sufficiently reactive toward protein carboxyl groups *in vacuo* at low LpH values to be a useful and versatile chemical modification reagent. The *in vacuo* nonaqueous methodology for chemically modifying protein has made it possible to examine functional group reactivities over a wide pH range by ^{13}C-NMR spectroscopy in order to obtain structural information. The objective of the study was to define the conditions for and to characterize the *in vacuo* reactions of ^{13}C-iodomethane and acidic ^{13}C-methanol (or ^{13}C-ethanol) with carboxyl groups in protein molecules, to characterize the different ^{13}C-methyl (or ^{13}C-ethyl) esters formed (i.e., α-, γ- and δ-carboxyl methyl esters of C-terminal, aspartyl, and glutamyl

Table 3-1. Experimental Model Systems to Investigate Protein Carboxyl Methylation[†]

Model System	Possible endogenous substrate	Postulated significance
Mammalian system		
Hypothalamus	Synaptosome membrane	Membrane depolarization Release of neurotransmitter
Posterior pituitary	Neurophysin	Exocytosis of vasopressin
Anterior pituitary	ACTH, LH, FSH, etc.	Hormonal action
Parotid gland	Vesicular membrane	Secretion of amylase
Adrenal gland	Chromaffin vesicle-membrane Chromogranin A	Secretion of catecholamine
Testis	Seminiferous tubule Spermatid	Function of pituitary-gonadal axis; Locomotion of sperm
Erythrocyte	Glycophorin A, band 4.5	Glucose transport
Platelet	14, 000 g pellets	Platelet activation
Neutrophil	Receptor for chemotactic peptides	Chemotaxis
Lymphocyte		Chemotaxis
Bacterial system		
Escherichia coli	Protein S3 and S9 Cytoplasmic membrane protein	Chemotaxis
Salmonella typhimurium	Cytoplasmic membrane protein	Chemotaxis

[†] Data taken from Paik and Kim, 1980 (c), and references therein.

amino acid residues, respectively), to determine the ^{13}C-chemical shifts for these derivatives, and to determine the extent of methyl- or ethyl-esterification of the model protein ribonuclease A. A comprehensive study of several proteins was carried out and the information observed is available to facilitate the characterization of proteins modified with iodomethane, acidic alcohol or other methylating reagents for the purpose of carrying out structure-function studies, and for C-terminal sequence identification, which is an ongoing field of research in our laboratory. The investigations of this study have provided further supporting evidence for the general applicability of the *in vacuo* nonaqueous methodology to effect the chemical modification of proteins.

2. Materials and Methods

2.1 Proteins

Ribonuclease A (bovine; Lot 127H1104), insulin (bovine; Lot 88H1307), α-chymotrypsin (bovine; Lot 91H7195), ovalbumin (grade V; Lot 106H7070), lysozyme (chicken egg-white, grade I; Lot 89F8275), and trypsin inhibitor (soybean, Type II-S; Lot 34H7050) were purchased from Sigma Chemical Company. Albumin (human; Lot M66206) was purchased from Armour Pharmaceutical Company. High purity insulin and mutant recombinant insulin [Lys(B28), Pro(B29)]; solution preparations intended for human injection were provided by Dr. Harold Rode of the Biologics and Genetic Therapies Directorate, Health Canada, and were either of bovine or pork origin from licensed manufacturers. *Bacillus thuringiensis* toxin was prepared from crystal protein via the published procedure of Bietlot *et al.* (1989).

2.2 Chemicals and Solvents

^{13}C-Iodomethane and ^{13}C-methanol were purchased from Sigma-Aldrich Chemical Company. ^{13}C-Ethanol was purchased from Cambridge Isotope Laboratories. Dipeptides (Leu-Leu, Gly-Phe, Phe-Gly, Val-Gly, Asp-Gly, Ala-Gly, Gly-Ala, Ala-Ala, Val-Ala, Arg-Asp, Ala-Asp, Asp-Ala, Glu-Ala) were purchased from Sigma-Aldrich Chemical Company. All other chemicals, reagents, and solvents were high-purity preparations obtained from commercial sources.

2.3 Preparation of N^{α}-Acetyl-Dipeptide ^{13}C-Methyl Ester Standards

Dipeptide (1 mg in 100 µL distilled water) was lyophilized in 0.6 mL eppendorf tubes from acidic buffer solution (pH 4.4) containing glacial acetic acid (2.7 µL), pyridine (2 µL), and additional distilled water (300 µL). The small eppendorf tube containing the dipeptide at LpH 4.4 was fitted into a large screw-cap Pyrex tube (16x200mm). The Pyrex tube was narrowed, reagent was added, and the tube was sealed according to the procedure of Taralp and Kaplan (1997). Dipeptide samples were incubated *in vacuo* (17 hr, 75 °C) using ^{13}C-methanol (10 µL) and 4 N HCl in dioxane (20 µL) to effect esterification. Sample tubes were cracked open and evacuated under vacuum for 30 min in the presence of NaOH pellets to remove excess reagent and catalyst. Esterified dipeptides were acetylated using acetic acid (200 µL) and acetic anhydride (100 µL). The extent of acetylation was monitored by the ninhydrin test and complete reaction occurred after 93 h at room temperature, as evidenced by the disappearance of colour with cadmium acetate/ninhydrin/acetone solution of samples spotted (10 µg) on Whatman 3 MM paper compared to controls (starting materials). Reagent and solvent were evacuated by speed vacuum. Samples were dissolved in 0.1% formic acid (500 µL) with vortex mixing and subsequently evacuated by speed

vacuum prior to analysis by ^{13}C-NMR spectroscopy. Dipeptide derivatives appeared as crystalline powders, dry powders, or opaque viscous gels.

While the above protocol was successful, the following ones were not so. Dipeptide, lyophilized directly in Pyrex tubes, was reacted *in vacuo* (51 h, 75°C) with ^{13}C-iodomethane. Acetylation was attempted with dry pyridine (200 µL) and acetic anhydride (50 µL), but most samples were poorly soluble in the pyridine. The portion of insoluble material was solubilized with acetic acid and acetic anhydride as above, but samples were reacted for 3 h and the acetylation reaction was incomplete. Nevertheless, the ^{13}C-NMR spectroscopic results were poor and peaks were on the order of those from natural abundance ^{13}C-atoms, indicating poor reaction with ^{13}C-iodomethane. To further verify the latter result, two dipeptides, Asp-Gly and Ala-Gly, lyophilized in eppendorf tubes, were reacted with ^{13}C-iodomethane *in vacuo* in Pyrex tubes (24 h, 100 °C), and the extent of reaction of the carboxyl groups was tested by high-voltage paper electrophoresis run at pH 4.4 at 3 kV for 30 min. A successful reaction would be evidenced by the absence of acidic peptides compared to control dipeptides. The results showed that the reaction did not proceed well with ^{13}C-iodomethane, as acidic bands were observed in both product samples and the corresponding starting materials (data not shown). In an alternative protocol, the same two dipeptides were first acetylated (16 h) with acetic acid/acetic anhydride (200 µL:100 µL) and then methylated with ^{13}C-iodomethane (10 µL). The results were fair in that similar chemical shifts for the ^{13}C-methyl ester derivatives of the dipeptides were observed as when prepared with ^{13}C-methanol/HCl; however, the extent of reaction with ^{13}C-iodomethane was much less than with the latter reaction, as evidenced by the much smaller peak intensities on the ^{13}C-NMR spectra.

2.3.1 Verification of N^{α}-Acetyl-Dipeptide ^{13}C-Methyl Ester Standards by Aqueous Reaction with Hydroxylamine

Samples of derivatized dipeptides (N^{α}-Ac-Ala-Ala-$O^{13}CH_3$, N^{α}-Ac-Ala-Asp(-$O^{13}CH_3$)-$O^{13}CH_3$, N^{α}-Ac-Asp(-$O^{13}CH_3$)-Ala-$O^{13}CH_3$), were split and half the sample was reacted with 2.5% hydroxylamine at pH 9 in D_2O with acetonitrile present as reference standard (see below). Samples were analyzed by ^{13}C-NMR spectroscopy.

2.4 *In Vacuo* Methylation of Proteins with ^{13}C-Iodomethane

Proteins (20 mg) were lyophilized without buffer from the following volumes in distilled water after adjustment of the appropriate solution pH with NaOH or HCl: soluble protein (1 mL); insulin (40 mL); and chicken egg-white lysozyme, LpH 10 (5 mL). Insulin (commercial insulin (Sigma), recombinant insulin and insulin mutant) and human albumin were dialyzed against 3x 4 L distilled water (3500 MWCO dialysis tubing) prior to lyophilization to remove excipients. Ribonuclease A, α-chymotrypsin, ovalbumin, and chicken egg-white lysozyme were lyophilized directly from solutions of the commercial preparations. Samples were flash frozen with liquid nitrogen to limit proteolysis. Lyophilized samples were desiccated prior to reaction to maintain protein dryness. *In vacuo* reactions with ^{13}C-iodomethane (25 μL) were carried out in sealed reaction vessels, which were placed in an oven at 75°C for 24 h according to the procedure described by Taralp and Kaplan (1997). After the reaction, excess reagent was recovered by trapping with liquid nitrogen as previously described (Taralp and Kaplan, 1997). Reaction vessels were cracked open and dried by extensive evacuation (typically 2-3 h) to remove any residual reagent and ^{13}C-methanol side-product.

Bt toxin (73 mg) was prepared with 200 μL 0.2 M sodium metaborate buffer, pH 10 and enough 1 N NaOH to adjust the solution pH to 10. The sample was lyophilized and reacted *in vacuo* with ^{13}C-iodomethane as described above. After reaction, as much sample as possible was dissolved in distilled water and insolubles were separated out using a microfuge; the supernatant was analyzed by ^{13}C-NMR spectroscopy (Varian 300 MHz spectrometer). After the first NMR spectrum was acquired, the sample was dialyzed against 8 M urea and then against distilled water, (prepared with 700 μL urea/D_2O, 50 μL phosphate buffer, pH 7.5, and 10 μL acetonitrile) lyophilized and re-analyzed by ^{13}C-NMR spectroscopy. The sample was subsequently dialyzed against distilled water after the second NMR spectrum was acquired, then lyophilized in a hydrolysis tube and hydrolyzed with 6 N HCl, and prepared for a third NMR analysis (Bruker 500 MHz spectrometer) with 800 μL D_2O, 30 μL acetonitrile adjusted to pH 7.5 with NaOH. The spectrum of the hydrolyzed sample is shown in Chapter 2 (Figure 2-9(b)).

2.4.1 LpH Course Experiments

Proteins (20 mg) were lyophilized without buffer from the appropriate afore-mentioned volumes at various LpH values: 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 8, 9 and 10. Insulin, an exception, was lyophilized from various volumes due to solubility limitations. The lyophilization volumes (Lvol) required for the preparation of insulin samples at the various LpH values were as follows, LpH (Lvol in mL): 2 (1), 3 (1), 3.5 (1), 4 (1), 4.5 (5), 5 (40), 5.5 (40), 6 (40), 6.5 (40), 7 (20), 8 (20), 9 (10), and 10 (10). However, it has been observed that some commercial preparations of insulin are more soluble at the lower pH values than others. For the latter cases, a larger Lvol is required.

Insulin samples were lyophilized from scintillation vials containing no more than 10 mL dissolved protein solution. For example, for LpH 5 insulin, four vials were used and the contents (lyophilized powders) of all four vials were combined and transferred into Pyrex tubes for reaction. The use of scintillation vials allows for a more quantitative transfer compared to retrieval from round or pointy bottom flasks, where static becomes a problem and sample retrieval is more difficult.

2.4.2 Time Course Experiments

Ribonuclease A and α-chymotrypsin were lyophilized from 1 mL total volume containing 20 mg/mL protein at pH 5. Samples were methylated *in vacuo* at 75°C with ^{13}C-iodomethane (25 μL) for various times (0, 15, 30 min, 1, 2, 4, 8, 12, 24, 48, and 72 h) according to the procedure described by Taralp and Kaplan (1997).

2.4.3 Effect of Buffers, Temperature, Counterions, and Reagent Amount

The effect of buffers on the reactivity of carboxyl groups with ^{13}C-iodomethane was examined. Ribonuclease A (20 mg) was dissolved in 0.8 mL distilled water and 0.2 mL 200 mM sodium phosphate buffer, pH 6, or 200 mM sodium metaborate buffer, pH 10, and the solution pH was adjusted to the appropriate value with 0.1 N or 1 N HCl and 0.1 N or 1 N NaOH prior to lyophilization. Samples were methylated *in vacuo* at 75°C or 100°C with ^{13}C-iodomethane (25 μL) for 24 and 72 h according to the procedure described by Taralp and Kaplan (1997).

The effect of counterions on the reactivity of carboxyl groups with ^{13}C-iodomethane was examined. Ribonuclease A (40 mg) was dialyzed against 3x 4 L distilled water (3500 MWCO dialysis tubing) from a 1 M LiCl, CsCl, or tetramethylammonium chloride (TMACl) aqueous solution of 5 mL. The control sample

of ribonuclease A was dialyzed without additional salt and was assumed to contain sodium as the counterion. Samples (10 mL) were split, their solution pH was adjusted to 6 with the appropriate hydroxide (0.1 or 1 N NaOH, LiOH, CsOH or TMAOH), and lyophilized in large screw-cap Pyrex tubes (16x200mm) from a total volume of approximately 5 mL (the increase in volume was due to washing of the dialysis tubing, ensuring quantitative transfer). Samples were methylated *in vacuo* at 75°C with ^{13}C-iodomethane (25 μL) for 24 h according to the procedure described by Taralp and Kaplan (1997).

The full extent of reaction of carboxyl groups with ^{13}C-iodomethane was verified by doubling the amount of reagent (50 μL) as compared to that typically used (25 μL) for the *in vacuo* methylation of ribonuclease A at LpH 6.

2.4.4 Consecutive *In Vacuo* Methylation Reactions

In order to test whether or not the liberation of HI from the methylation reaction with iodomethane inhibits the extent of reaction of carboxyl groups, and to test the hypothesis that surface exposed carboxyl groups react more readily than buried carboxyl groups, the following consecutive *in vacuo* methylation reactions were carried out. Two samples of ribonuclease A (20 mg each) were lyophilized at LpH 6 from a 1 mL volume and reacted *in vacuo* at 75°C with iodomethane (25 μL) for 24 h according to the procedure described by Taralp and Kaplan (1997). A consecutive *in vacuo* methylation reaction was carried out on one of these unlabeled methylated samples using ^{13}C-iodomethane (25 μL) for an additional 24 h at 75°C; this was the control sample. The other unlabeled methylated sample was re-dissolved in distilled water and its solution pH was adjusted from 2.5 to 6. The sample was re-lyophilized and re-reacted *in vacuo* at

75°C with ^{13}C-iodomethane (25 μL) for 24 h. Similarly, with two additional ribonuclease A samples, reactions were carried out consecutively *in vacuo* using ^{13}C-iodomethane (25 μL) for both steps. These samples were positive, maximum reaction controls. Consecutive reaction samples were compared to ^{13}C-methylated ribonuclease A reacted for 24 h at 75°C.

Alternatively, ribonuclease A (40 mg) was lyophilized from a 2 mL volume (20 mg/mL protein) at LpH 6 and reacted *in vacuo* at 75°C with iodomethane (50 μL). The methylated sample was split by mass and consecutive *in vacuo* reactions were carried out as described above using ^{13}C-iodomethane (25 μL) without (control) and with re-lyophilization at LpH 6 in between *in vacuo* reactions.

Moreover, ribonuclease A (30 mg) was lyophilized from a 1.5 mL volume (20 mg/mL protein) at LpH 6 and reacted *in vacuo* at 75°C with iodomethane (40 μL). The methylated sample was split by mass into three 10 mg portions and consecutive *in vacuo* reactions were carried out as described above on two of the portions using ^{13}C-iodomethane (12.5 μL) without (control) and with re-lyophilization at LpH 6 in between *in vacuo* reactions. A similar protocol was performed for reactions using ^{13}C-iodomethane for initial (24 h) and consecutive reactions (2x24 h).

2.5 *In Vacuo* Esterification of Proteins with ^{13}C-Methanol or ^{13}C-Ethanol and Hydrochloric Acid Catalyst

Proteins (20 mg) were dialyzed (against 3x 4 L distilled water) from a 200 mM NaCl aqueous solution; the solution pH was adjusted to 3 and samples were lyophilized without buffer from the following volumes: soluble ribonuclease A or α-chymotrypsin (1 mL), insulin (10-15 mL). Soluble protein samples were lyophilized directly in small culture tubes, which were subsequently fitted into 16x150mm

borosilicate glass culture tubes. Insulin was lyophilized in scintillation vials and then transferred to the small culture tubes; the latter were then placed into the larger culture tubes. Lyophilized samples were desiccated prior to reaction to maintain protein dryness. Samples were esterified *in vacuo* at 75°C with ^{13}C-methanol (25 μL) or ^{13}C-ethanol (30 μL) and 4 N HCl in dioxane (10 μL) for 24 h according to the methodology described by Taralp and Kaplan (1997). To our knowledge, this is the first time an esterification has been carried out on protein in the gas phase.

2.5.1 Time Course Experiments

Dialyzed ribonuclease A (20 mg) was lyophilized at LpH 3 from 1 mL total volume in a small culture tube, which was fitted into a 16x150mm borosilicate glass culture tube. Samples were esterified *in vacuo* at 75 °C with ^{13}C-methanol (25 μL) and 4 N HCl in dioxane (10 μL) for various times (0, 15, 30 min, 1, 2, 4, 8, 12, 24, 48, and 72 h) according to the methodology described by Taralp and Kaplan (1997). Samples were esterified *in vacuo* at 75°C with ^{13}C-ethanol (30 μL) and 4 N HCl in dioxane (10 μL) for 24 and 48 h.

2.5.2 Consecutive *In Vacuo* Esterification Reactions

In order to verify the extent of reaction of carboxyl groups with methanol and acid catalyst, and to test the hypothesis that surface exposed carboxyl groups react more readily than buried carboxyl groups, the following consecutive *in vacuo* esterification reactions were carried out. Two samples of ribonuclease A (20 mg each) were lyophilized at LpH 3 from a 1 mL volume and reacted *in vacuo* at 75°C with anhydrous methanol (25 μL) and 4 N HCl in dioxane (10 μL) for 24 h according to the methodology described by Taralp and Kaplan (1997). A consecutive *in vacuo* esterification reaction

was carried out on one of these unlabeled esterified samples using ^{13}C-methanol (25 μL) and 4 N HCl in dioxane (10 μL) for an additional 24 h at 75°C; this was the control sample. The other unlabeled esterified sample was re-dissolved in distilled water and its solution pH was re-adjusted to 3. The sample was re-lyophilized and re-reacted *in vacuo* at 75°C with ^{13}C-methanol (25 μL) and 4 N HCl in dioxane (10 μL) for 24 h. Similarly, with two additional ribonuclease A samples, reactions were carried out consecutively *in vacuo* using ^{13}C-methanol (25 μL) and 4 N HCl in dioxane (10 μL) for both steps. These samples were positive, maximum reaction controls. Consecutive reaction samples were compared to ^{13}C-methyl-esterified ribonuclease A reacted for 24 h.

Alternatively, ribonuclease A (20 mg) was lyophilized from a 1 mL volume at LpH 3 and reacted *in vacuo* at 75°C for 24 h with ^{13}C-methanol (25 μL) and 4 N HCl in dioxane (10 μL). The esterified sample was split by mass into two 10 mg portions and consecutive *in vacuo* esterification reactions were carried out as described above using ^{13}C-methanol (12.5 μL) and 4 N HCl in dioxane (5 μL) without (control) and with re-lyophilization at LpH 3 in between *in vacuo* reactions.

Moreover, ribonuclease A (30 mg) was lyophilized from a 1.5 mL volume (20 mg/mL protein) at LpH 3 and reacted *in vacuo* at 75°C with ^{13}C-methanol (38 μL) and 4 N HCl in dioxane (10 μL). The esterified protein sample was split by mass into three 10 mg portions and consecutive *in vacuo* reactions were carried out as described above on two of the portions using ^{13}C-methanol (12.5 μL) and 4 N HCl in dioxane (5 μL) without (control) and with re-lyophilization at LpH 3 in between *in vacuo* reactions; the third 10 mg portion was analyzed to determine the extent of reaction after one incubation period (24 h, 75°C). Additionally, three lyophilized ribonuclease A

(30 mg) samples at LpH 3 were reacted *in vacuo* at 75°C with ^{13}C-methanol (38 µL) and 4 N HCl in dioxane (10 µL). The samples were split into 10 mg and 20 mg portions by mass. The 10 mg samples, as well as one of the 20 mg samples, were used as initial reaction controls. Consecutive reactions were carried out on the two 20 mg samples without and with re-lyophilization at LpH 3 in between *in vacuo* reactions. This latter protocol was the best approach, since the extent of the initial reactions (24 h, 75°C) could be assessed.

2.6 Identification of ^{13}C-Methyl Ester Resonances: Reaction of ^{13}C-Methyl Esters with Hydroxylamine to Form Hydroxamates and ^{13}C-Methanol

After initial ^{13}C-NMR spectroscopic analysis, the ^{13}C-methyl-esterified protein was reacted with 2.5% hydroxylamine at pH 9 (Darbre, 1986; Halpin and Richardson, 1985). Samples were heated at 78°C for 30 minutes in screw-cap NMR tubes (5 mm; Wilmad Glass Co.), then either analyzed directly or frozen at -20°C until they were analyzed via ^{13}C-NMR spectroscopy. If necessary, samples were treated further with base, by adjusting to pH 11.5 with 1 N NaOH, and heated for an additional 30 minutes at 78°C, followed by a readjustment of the pH to 7.5 with 1 N HCl, and subsequently analyzed by ^{13}C-NMR spectroscopy.

2.6.1 *In Vacuo* Reaction of ^{13}C-Esterified Protein with Hydroxylamine to Form Hydroxamates with Carboxyl Esters and ^{13}C-Alcohol

For some proteins (e.g., α-chymotrypsin, ovalbumin, human albumin), the hydroxylamine reaction in aqueous medium leads to a gel state, which is indicative of an intermediate denaturation state (Magdassi, 1996; Ferry, 1948). Since reactions were carried out in screw-cap NMR tubes, the gel state observed when the protein was treated with hydroxylamine and base did not allow for complete conversion of methyl esters to

hydroxamates for the purpose of identifying the ^{13}C-resonances that arose from ^{13}C-methyl esters in the region of the ^{13}C-NMR spectra between 50-57 ppm. The resultant gel state also led to the need for overnight NMR acquisitions in order to obtain a good quality spectrum with a good signal-to-noise ratio. In order to circumvent this problem, the hydroxylamine reaction was carried out *in vacuo* on the dried derivatized protein [^{13}C-methylated (with ^{13}C-iodomethane) or ^{13}C-esterified ribonuclease A (with ^{13}C-alcohol and HCl catalyst); 10-20 mg] at 75°C with hydroxylamine (30 μL 50 wt.% hydroxylamine in water, pH 9) at various times (0, 15, 30 min, 1, 2, 4, 8, 12, and 24 h). After reaction for up to 2 h the protein was in an insoluble gel state (i.e., the physical state of the protein was a gel when the sample was rehydrated in D_2O, acetonitrile and 3-pentanol internal standard), and thus, inconvenient form for ^{13}C-NMR spectroscopic analysis. The 4 h time point gave the optimum result in that the protein was soluble under the conditions used for ^{13}C-NMR spectroscopic analysis; and, it was sufficient to remove all resonances due to ^{13}C-methyl esters from the spectra of ^{13}C-methylated ribonuclease A without significant degradation of the protein, which occurred with longer reaction times. Thus, although it took a longer time (4 h at 75°C) to achieve complete hydroxamate derivatization of carboxyl methyl esters compared to the aqueous protocol, the *in vacuo* procedure was found to be better than the latter for our purposes because it was possible to solubilize the problematic proteins. The latter proteins by-passed the gel state denaturation phase, which also was evident in the *in vacuo* protocol up to 2 hours of heating time. The four hours of heating time is the optimal time in terms of solubility and also in terms of exhaustive degradation of observable trimethylamino resonances.

2.6.2 Aqueous Reaction of ^{13}C-Methylated Ribonuclease A with Hydroxylamine and Sodium Hydroxide: A ^{13}C-NMR Spectroscopic Kinetic Study for the Identification of ^{13}C-Resonances Arising from ^{13}C-Methyl Esters

^{13}C-Methylated ribonuclease A (LpH 5, 48 h, 75°C) was analyzed by ^{13}C-NMR spectroscopy in 8 M urea/D_2O in the presence of acetonitrile (4 spectra acquired). After acquisition of the initial spectrum (spectra), the sample was reacted with 2.5% hydroxylamine (30 µL 50 wt.% hydroxylamine in water, pH 9) at room temperature (9 spectra acquired at pD 8.7) and the disappearance of the resonances corresponding to the ^{13}C-methyl esters of protein carboxyl groups was observed at specific time blocks over the course of a multi-hour acquisition. The sample was stored at −10°C overnight in between acquisitions when required. The pD of the aqueous reaction with hydroxylamine was initially at 8.7, and then increased to pD 10.0 (5 spectra acquired). The reaction was carried out at room temperature for 7.5 h and the sample was heated at 75°C for two times 30 min and then for 2 h (3 spectra acquired). In order to find the end-point of the experiment, the pD of the solution was increased to 12.0 and the sample was heated for 1 h at 75°C. The sample was then dialyzed against 3x 4 L distilled water (3500 MWCO dialysis tubing), speed vacuumed, and a final spectrum was acquired at pD 5.4. Another sample of ^{13}C-methylated ribonuclease A (LpH 5, 72 h, 75°C) was used to study the hydroxylamine reaction. The sample pD was adjusted to 9.9 after the addition of hydroxylamine (30 µL 50 wt.% hydroxylamine in water, pH 9) and the sample was heated at 75°C for two 30 min intervals.

2.7 ^{13}C-NMR Spectroscopic Analysis of ^{13}C-Esterified Proteins

^{13}C-NMR spectra were obtained using a Bruker spectrometer operating at 9.4 Tesla (^{1}H, 400.13 MHz; ^{13}C, 100.6 MHz). ^{13}C-Methylated protein was dissolved in deuterium oxide containing 8 M ultra pure urea (700 µL; pH 2.5; Schwarz/Mann Biotech;

Lot 4116). Acetonitrile (30 μL) was added to the protein sample as an internal chemical shift reference (Breitmaier, E., and Voelter, W., 1987a) ($^{13}CH_3$, +1.70 ppm). For quantification, 3-pentanol (40 μL) ($^{13}CH_3$, 10.14 ppm; $^{13}CH_2$, 29.44 ppm; ^{13}CH, 75.41 ppm) was included (Breitmaier, E., and Voelter, W., 1987b). The sample solution was adjusted to pD 5.4 [pD = pH meter reading + 0.4; Glasoe and Long, 1960] using 0.1 N or 1 N HCl and/or NaOH, transferred to an NMR tube (5 mm), and analyzed. Spectra were acquired typically for 600 scans using the DEPT-135 pulse sequence (Bendall *et al.*, 1982). For the hydroxylamine and base-treated samples, the ^{13}C-NMR spectra were acquired at pD 9.4 and 7.9, respectively. It is noteworthy that for hydroxylamine and base-treated α-chymotrypsin, the spectra were acquired overnight (18000 scans) to improve the signal-to-noise ratio, as the protein was observed to be in a gel state after these treatments. Samples of ^{13}C-methylated proteins that formed gels with hydroxylamine and base treatment in the aqueous protocol (e.g., α-chymotrypsin, ovalbumin, human albumin) were reacted with hydroxylamine *in vacuo* (see Section 2.6.1 above) and spectra were acquired for 600 scans.

For LpH course and time course experiments, all ^{13}C-methylated proteins were dissolved in 8 M urea/D_2O for ^{13}C-NMR spectroscopic analysis. For further studies with ribonuclease A (e.g., effect of buffers, temperature, reagent amount, counterion, consecutive *in vacuo* experiments, repeated 24 and 72 h experiments, and esterification reactions with ^{13}C-methanol/HCl), the sample was dissolved in D_2O or acidic (pH 2.5) D_2O, followed by addition of acetonitrile and 3-pentanol (see above). Esterified insulin and α-chymotrypsin were solubilized in 8 M urea/D_2O, pH 2.5.

2.8 ^{13}C-NMR Spectroscopic Analysis of N$^{\alpha}$-Acetyl-Dipeptide ^{13}C-Methyl Ester Standards

^{13}C-NMR spectra were obtained using a Bruker spectrometer operating at 9.4 Tesla (^{1}H, 400.13 MHz; ^{13}C, 100.6 MHz). ^{13}C-Methylated acetylated dipeptide standard (0.5 or 1 mg) was dissolved in deuterium oxide (700 µL), followed by the addition of acetonitrile (25 µL), which was included as an internal chemical shift reference (Breitmaier and Voelter, 1987a). Sample "pH" or pD was adjusted to 5.4 with 1 N NaOH and/or 1 N HCl with vortex mixing. Spectra were acquired typically for 600 scans using the DEPT-135 pulse sequence (Bendall *et al*, 1982).

2.9 NMR Data Processing

NMR time domain data was Fourier transformed into frequency domain data using the WIN_1D NMR computer program. The peak height of the internal standard, 3-pentanol (75.4 ppm), was set to 10 cm and the peaks of interest were compared relatively. The internal standard peak (75.4 ppm) was integrated and set to an area of 100 for a semi-quantitative comparison of relative peak areas for peaks of interest between spectra. Spectra were saved as Windows meta files and processed further in Corel Draw 9 and Microsoft Word 98 (Windows 2000).

2.10 Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) of Esterified Ribonuclease A

Ribonuclease A and the esterified derivatives (methyl and ethyl esters of protein carboxyl groups; see Section 2.5 and Figure 3-22) were characterized by SDS-PAGE. The electrophoresis was carried out using the Pharmacia PhastGel™ system. PhastGels™ (20% acrylamide-bisacrylamide) were used. Sample buffers were 5x concentrated and were prepared in the presence and absence of 2-mercaptoethanol. The

recipe used was a modified version of the original recipe from Bollag and Edelstein (1991) and is as follows (10 mL total volume, pH 3): 5% (v/v) acetic acid (500 µL), 25% glycerol (2.5 mL), 5% SDS (1.54 mL 32.5% stock), +/- 0.72 M 2-mercaptoethanol (0.5 mL), 0.1% bromophenol blue (1 mL 1% stock), 4.36-4.86 mL distilled water. Samples were not heated before analysis. Gels were stained with Coomassie blue.

2.11 Estimation of the Extent of Modification of Carboxyl Groups of Methylated or Esterified Ribonuclease A via Difference Titration Methodology

An estimation of the number of carboxyl groups esterified in *in vacuo* methylated ribonuclease A preparations at various LpH values (LpH 3, 4, 5, 6, 7, and 10) was obtained by measuring the decrease in the number of ionizable groups that titrate between pH 3.0 and 6.0. The extent of esterification via the *in vacuo* methanol/HCl reaction (LpH 3) was determined in the same manner.

^{13}C-Methylated ribonuclease A (with ^{13}C-iodomethane), ^{13}C-methyl or ^{13}C-ethyl-esterified ribonuclease A (with ^{13}C-methanol/HCl) and native ribonuclease A (20 mg) samples were dissolved in 10 mL 0.01 M HCl (protein solution pH 2.1-2.5). Each sample was divided in half and the pH of the solutions was adjusted to between 2.75-2.8 with 1N NaOH, dispensed by a plastic syringe fitted with 1 mm Tygon tubing. Solutions were equilibrated under nitrogen for 10-15 minutes prior to commencing titration at room temperature. Samples were forward and backward titrated between pH 2.8 and 6 with 0.05 N NaOH or 0.05 N HCl, respectively. Base or acid was dispensed using an Agla syringe fitted with 1 mm Tygon tubing, connected to a plotter and gear system yielding 1 µL/mm. The total volume of base or acid used in the forward and backward titrations was determined by measuring the distance on the plotter paper between pH 3.0 and 5.5.

To verify the extent of reaction derivatized ribonuclease A (20 mg) samples were split and half the sample was analyzed by ^{13}C-NMR spectroscopy (see Figure 3-11), and the other half was titrated.

2.12 Estimation of the Extent of Modification of Carboxyl Groups of Methylated Ribonuclease A via High-Voltage Paper Electrophoresis

In an alternative, more visually qualitative and semi-quantitative method, the extent of esterification of ribonuclease A methylated at LpH 3, 6 and 10 with iodomethane was estimated from the electrophoretic mobility of peptic peptides of native and derivatized protein at pH 6.5. High-voltage paper electrophoresis was carried out at 2.5 kV/cm for 30 min. Peptides were detected with ninhydrin/cadmium acetate/acetone reagent.

2.12.1 Pepsin Digest of Native and Methylated Ribonuclease A

Native (protein blank) and methylated ribonuclease A samples (10 mg) that had been titrated for the determination of the extent of reaction of carboxyl groups with iodomethane were dialyzed against 0.01 M HCl and lyophilized in scintillation vials. Samples were dissolved in 98-100% formic acid, followed by the addition of 4.5 mL distilled water. Pepsin solution (1:20 pepsin:protein) in 10% formic acid was added (80 μL of 6.35 mg/mL 10% formic acid solution of pepsin were added), and samples were shaken at 37°C overnight. Samples were diluted 2-fold with distilled water and freeze-dried, followed by careful and quantitative transfer to eppendorf tubes and re-freeze-dried in the presence of P_2O_5 and NaOH pellets.

3. Results

3.1 Aqueous Reaction of ^{13}C-Methylated Ribonuclease A with Hydroxylamine and Sodium Hydroxide: A ^{13}C-NMR Spectroscopic Kinetic Study for the Identification of ^{13}C-Resonances Arising from ^{13}C-Methyl Esters

A ^{13}C-NMR spectroscopic kinetic study was carried out with ^{13}C-methylated ribonuclease A (LpH 5, 48 h, 75°C) for the identification of the ^{13}C-resonances that arise from ^{13}C-methyl esters. Removal of the ^{13}C-methyl esters arising from Asp and Glu residues (53.46 and 53.18 ppm, respectively; see below) was slow in the presence of hydroxylamine at pD 8.7; there was a slow but progressive increase in the resonance corresponding to ^{13}C-methanol. An increase in the solution pD to 10.0 increased the conversion of methyl esters to the hydroxamate derivatives, or increased the rate of hydrolysis by hydroxide anion; however, the resonance for the γ-carboxyl methyl esters of Asp residues did not diminish completely (after 6 h in the presence of hydroxylamine at room temperature), even after the sample was heated for a total of 3 h. (The broadened resonance corresponding to the methyl esters of the Asp residues could be qualitatively observed more easily than that of the Glu residues because the latter overlaps the resonance for the trimethyl α-amino Lys.) The resonance for the γ-carboxyl methyl esters of Asp residues only diminished significantly when the sample pD was increased to 12.0 and the sample was heated for 1 h at 75°C. The sample was dialyzed and the resonances remaining were those arising from the trimethyl ε-amino group of lysyl residues, the trimethyl α-amino group of the N-terminal lysyl residue, and the trimethylamino group derived from an impurity at 54.15 ppm. A separate sample of ^{13}C-methylated ribonuclease A (LpH 5, 72 h, 75°C) was used to study the hydroxylamine reaction. The sample pD was adjusted to 9.9 after the addition of hydroxylamine and the sample was heated at 75°C for two 30 min intervals. The resonance for the methyl esters

of the Asp residues decreased substantially, but again, the result was an incomplete conversion to the hydroxamate groups. Thus, it was still necessary to increase the pD to 11.9 and heat for an additional hour or longer in some cases. These results led to the trial of an *in vacuo* hydroxylamine reaction (see Section 2.6.2, Chapter 3; Sections 2.5.1, 2.5.2, 3.1 and 4, Chapter 2).

Thus, in some cases, removal of the methyl ester resonances from the ^{13}C-NMR spectrum was not complete with either hydroxylamine or base treatment or both in aqueous medium. In other cases, such as was the case with α-chymotrypsin, ovalbumin and human albumin, the proteins gelled after hydroxylamine treatment and the conversion of methyl esters to hydroxamates was incomplete. For these reasons, an *in vacuo* hydroxylamine reaction (30 μL aliquot of 50 wt.% hydroxylamine in water, pH 9) was carried out on ribonuclease A. The removal of the methyl ester resonances was achieved with an optimal 4 h reaction time with respect to protein solubility for the NMR spectroscopic analysis (see Section 2.6.1). If maintainance of structural integrity and activity is desired, for problematic proteins that gel with the aqueous hydroxylamine reaction, a mild saponification reaction may be carried out instead with a larger volume at a lower temperature (i.e., 25 or 37°C) for a longer duration of time (Broomfield *et al.*, 1965; Acharya and Vithayathil, 1975).

3.2 *In Vacuo* Methylation of Model Proteins with ^{13}C-Iodomethane

The results obtained with several test proteins demonstrate that by controlling the pH of lyophilization (LpH) it is possible to selectively or preferentially methylate the carboxyl groups using ^{13}C-iodomethane. The α-amino group also reacts to a small extent to form the trimethylated amino acid residue at the lower LpH values (less than LpH 4).

Figures 3-1 to 3-6 illustrate the ^{13}C-NMR spectra of ^{13}C-methylated insulin, ribonuclease A, human albumin, α-chymotrypsin, ovalbumin and lysozyme, respectively, obtained by reaction of each protein with ^{13}C-iodomethane at various LpH values (LpH 3-10). Figure 3-7 illustrates the ^{13}C-NMR spectra of ^{13}C-methylated *Bacillus thuringiensis* toxin before and after dialysis against 8 M urea.

^{13}C-Resonances that arise from ^{13}C-methyl esters of carboxyl groups are observable in the region from 52-57 ppm. Assignment of the resonances arising from the ^{13}C-methyl esters of carboxyl groups was made on the basis of the chemical shifts determined with standard model compounds, namely acetylated ^{13}C-methyl-esterified dipeptides. Table 3-2 gives the chemical shifts of the methyl group in α-, γ-, and δ-carboxyl ^{13}C-methyl esters in N^{α}-acetylated dipeptides. The chemical shift of the γ-carboxyl ester of the aspartyl side-chain (53.43 ppm) does not vary significantly with the composition of the dipeptide and is well separated from that of the δ-carboxyl side-chain of glutamic acid (53.10 ppm). In contrast, the chemical shifts of the α-carboxyl ^{13}C-methyl esters vary with the terminal amino acid and composition of the peptide and are always downfield from the resonances of the γ- and δ-carboxyl ^{13}C-methyl esters. When the modified protein was treated with base or hydroxylamine (a carboxyl methyl ester group-specific reagent), all the resonances disappeared with the concomitant appearance of the resonance for ^{13}C-methanol at 49.8 ppm, showing that they arose from esterification of the various classes of carboxyl groups in the protein.

As previously demonstrated (Taralp, 1997, Doctoral Thesis; Chapter 2 herein, Vakos *et al.*, 2000), resonances arising from trimethylamino groups also reside in this spectral region. Chemical shifts of other potentially reactive protein functional groups

were assigned by referencing the Table of ^{13}C-methylated amino acid and peptide standard compounds prepared by Taralp (Doctoral Thesis, 1997a).

Time course experiments were carried out for esterified ribonuclease A (with ^{13}C-iodomethane and ^{13}C-methanol/HCl) and α-chymotrypsin (with ^{13}C-iodomethane), and the results are presented below for the respective proteins.

3.2.1 Insulin

The stability of carboxyl ^{13}C-methyl esters in methylated insulin was studied. One of the α-carboxyl ^{13}C-methyl ester groups (54.00 ppm; Asn C-terminus on A-chain) was progressively hydrolyzed with time at 300 K, with the concomitant increase in the resonance for ^{13}C-methanol at 49.8 ppm (3.6, 4.5, 5.2, 77% increase at 30 min, 1, 1.5, and 30 h, respectively; complete removal of the α-carboxyl ^{13}C-methyl ester peak at 54.00 ppm was observed when the sample was re-analyzed by ^{13}C-NMR spectroscopy after 7 weeks at 4°C). The resonance at 54.00 ppm corresponds to a carboxyl methyl ester that is more unstable or labile than the other carboxyl methyl esters, and is likely due to the tendency of the asparaginyl side-chain to cyclize and form a substituted succinimide that is susceptible to nucleophilic attack (i.e., by water). The ester groups in methylated proteins were progressively hydrolyzed with increasing pD, and a higher temperature accentuated this effect (see Section 3.1).

Figure 3-1 shows the ^{13}C-NMR spectra of ^{13}C-methylated insulin (from Sigma Chemical Company), obtained by reaction with ^{13}C-iodomethane at LpH values between 3 and 10. Figures 3-1a and 3-1b demonstrate the preferential methylation of the two α- and four δ-carboxyl groups of insulin reacted at LpH 3 and LpH 3.5. The sharp ^{13}C-resonances at 54.00 and 53.76 ppm correspond to the two α-carboxyl methyl esters

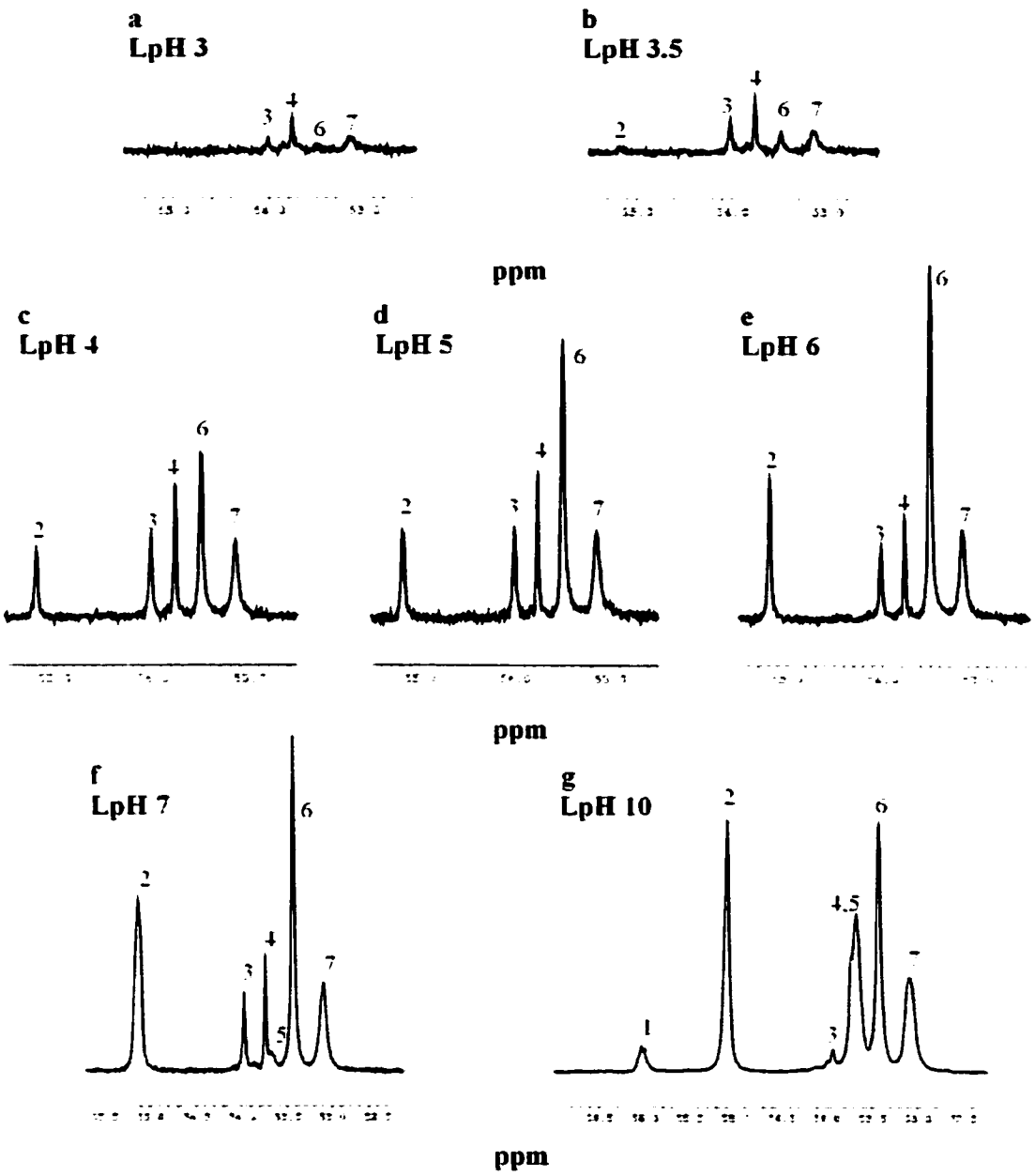


Figure 3-1. Expanded ^{13}C NMR spectra of ^{13}C-methylated insulin at various LpH values (LpH 3-10; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) Tyr-O**Me** (56.04ppm); (2) **Me**$_3$N$^{\alpha+}$-Gly (55.18 ppm); (3) C$_\alpha$OO**Me** (A-chain Asn. 54.00 ppm); (4) C$_\alpha$OO**Me** (B-chain Ala. 53.76 ppm); (5) **Me**$_3$N$^{\varepsilon+}$-Lys (53.65 ppm); (6) **Me**$_3$N$^{\alpha+}$-Phe (53.49 ppm); (7) C$_\delta$OO**Me** (four overlapping Glu residues. 53.15 ppm).

and are downfield from those of the four δ-carboxyl methyl esters; the latter resonances overlap each other to form a broad peak at 53.15 ppm. The two α-carboxyl methyl ester resonances correspond to the C-terminal Asn and Ala residues of the A and B chains of insulin, respectively. As the LpH is raised, the intensity of these peaks increases along with the appearance of the resonances of the trimethylated α- and ε-amino groups, which are dominant at alkaline pH values. It is particularly noteworthy that the ^{13}C-NMR spectra of two ^{13}C-methylated insulin samples prepared from high purity human insulin and a mutant recombinant insulin {[Lys(B28), Pro(B29)] intended for human injection} were similar to those of the insulin purchased from Sigma, except for the observation that the ^{13}C-methylated mutant recombinant insulin sample, which has the B28 and B29 amino acid residues switched, gave a resonance for one of its C-terminal carboxyl methyl esters (B-chain Ala) at 53.83 ppm, as opposed to the 53.75 ppm chemical shift observed for the normal ^{13}C-methylated human insulin sample, which was similar to that observed for the ^{13}C-methylated insulin purchased from Sigma, as expected (see Table 3-3). Clearly in the insulin mutant sample, the chemical shift of the C-terminal Ala is shifted downfield as a result of the penultimate amino acid residue being Pro instead of Lys. The chemical shifts of the α-carboxyl ^{13}C-methyl ester of the B-chain carboxyl terminal Ala for human insulin (-Pro_{28}-Lys_{29}-Ala_{30}-O$^{13}CH_3$) and that of a recombinant mutant (-Lys_{28}-Pro_{29}-Ala_{30}-O$^{13}CH_3$) differ by 0.08 ppm.

A small proportion of the N^{α}-trimethylated Phe resonance is observed at 53.49 ppm at LpH 3 and LpH 3.5, and the N^{α}-trimethylated Gly resonance becomes evident at LpH 3.5 at 55.18 ppm. It is noteworthy that insulin has one lysyl residue, which is observed as its N^{ε}-trimethylated form at LpH 7 or higher. Its broad

^{13}C-resonance at 53.65 ppm overlaps (slightly upfield) with one of the resonances corresponding to an α-carboxyl ^{13}C-methyl ester, namely that of the C-terminal Ala residue of the B-chain of insulin (53.76 ppm). The broad peak of overlapped resonances corresponding to N^{ε}-trimethylated Lys residues overlap one of the C-terminal peaks (Ala of B-chain), but at LpH 10, the peak for the other C-terminus (Asn of A-chain) is diminished due to alkaline hydrolysis during sample preparation (Figure 3-1). It is noteworthy that the hydroxyl group of the phenolic side-chain of tyrosyl residues reacts in insulin *in vacuo* at LpH 10 (observed at 56.04 ppm on the ^{13}C-NMR spectrum), which suggests that this group is exposed to reagent. Since tyrosyl residues are typically buried in a protein molecule, the small insulin molecule is likely partially unfolded at LpH 10. Tyrosine is not observed to react *in vacuo* at LpH 10 with any of the other test proteins, suggesting that these proteins contain buried phenolic side-chains of tyrosyl residues.
In addition to dipeptide standard compounds used to assign the peak resonances of the δ-carboxyl ^{13}C-methyl esters, advantage was taken of the fact that insulin has four Glu residues and no Asp residues. The chemical shift of the resonances for the δ-carboxyl ^{13}C-methyl esters of Glu residues in insulin was used to verify the assignment of these groups in the other test proteins. The resonances observed for the α-carboxyl ^{13}C-methyl esters of methylated insulin were sharp compared to those of the other test proteins. The broadened α-carboxyl ^{13}C-methyl ester resonances observed with test proteins other than insulin is likely due to the presence of minor impurities resulting from proteolysis (especially with α-chymotrypsin). The latter would reveal ^{13}C-resonances in the C-terminal chemical shift range, which may overlap with the α-carboxyl ^{13}C-methyl ester

Table 3-2. ^{13}C-NMR Spectroscopic Chemical Shifts of *In Vacuo* Esterified N^{α}-Acetylated Dipeptide Standard Compounds

Ac-a.a.$_1$-a.a.$_2$-O^{13}CH$_3$	Minor Resonances (ppm)	($C_{\alpha}OO^{13}CH_3$) (ppm)	($C_{\gamma,\delta}OO^{13}CH_3$) (ppm)
Ac-Leu-Leu-O^{13}Me		53.65 (α)	
Ac-Gly-Phe-O^{13}Me	53.71, 53.62	53.79 (α)	
Ac-Phe-Gly-O^{13}Me	53.71	53.63 (α)	
Ac-Val-Gly-O^{13}Me		53.56 (α)	
Ac-Asp(O^{13}Me)-Gly-O^{13}Me	54.07 [60% 53.64 ppm (α)]	53.64 (α)	53.43 (γ)
Ac-Ala-Gly-O^{13}Me	53.75 [33% 53.62 ppm (α)]	53.62 (α)	
Ac-Gly-Ala-O^{13}Me		53.79 (α)	
Ac-Ala-Ala-O^{13}Me		53.74 (α)	
Ac-Val-Ala-O^{13}Me		53.65 (α)	
Ac-Arg-Asp(O^{13}Me)-O^{13}Me	53.86	54.05 (α)	53.43 (γ)
Ac-Ala-Asp(O^{13}Me)-O^{13}Me	53.74, 53.88 [20% 54.06 ppm (α)]	54.06 (α)	53.43 (γ)
Ac-Asp(O^{13}Me)-Ala-O^{13}Me		53.78 (α)	53.42 (γ)
Ac-Glu(O^{13}Me)-Ala-O^{13}Me		53.78 (α)	53.10 (δ)

resonances of the protein being studied at low LpH values. Significantly, trimethylamino resonances from such impurities must be evidenced at higher LpH values.

3.2.2 Ribonuclease A

Figure 3-2 shows the ^{13}C-NMR spectra of ^{13}C-methylated ribonuclease A, obtained by reaction with ^{13}C-iodomethane at LpH values between 3 and 10. Figure 3-2 demonstrates the preferential methylation of the α-, γ-, and δ-carboxyl groups of ribonuclease A reacted between LpH 3 and 5. The ^{13}C-resonances corresponding to the carboxyl methyl esters are: 53.71 (α), 53.56 (γ), 53.48 (γ), and 53.24 ppm (δ). There are two unidentified peaks, a small one at 54.23 ppm and a large sharp one at 56.21 ppm (not shown). The latter peak is believed to arise from the methylation of a small molecule.[1] The 56.21 ppm resonance disappeared after the methylated ribonuclease A sample was

---

[1]A sharp peak usually arises from a small molecule in a non-viscous solvent. In the "extreme narrowing" condition, where $\omega_o^2\tau_c^2<<1$, the longitudinal relaxation time, T_1, is equal generally to the transverse relaxation time, T_2, and linewidths are narrow ($\Delta\nu_{1/2}=1/\pi T_2^*$), where T_1^{-1} is proportional to τ_c. Proteins are macromolecules and have slower molecular motions compared to small molecules, and hence have different relaxation times ($T_2 \neq T_1$), relaxation rates ($T_2^{-1} \neq T_1^{-1}$), and spectral density functions (at the "non-extreme narrowing" condition $\omega_o^2\tau_c^2>>1$. T_2^{-1} is proportional to τ_c; while T_1^{-1} is proportional to τ_c^{-1}). The transverse relaxation (T_2), which affects linewidth, becomes increasingly efficient as correlation times (τ_c) (and hence, molecular weights) increase. Hence, a carbon-13 atom in a macromolecular protein will have a broader resonance than that in a small molecule in a non-viscous solvent ($\Delta\nu_{1/2}=1/\pi T_2^*$ and $\Delta\nu_{1/2}$ is proportional to τ_c).

In the "extreme narrowing" condition, where $\omega_o^2\tau_c^2<<1$, where ω_o is the Larmor frequency and τ_c is the correlation time, which is a time constant that characterizes the auto-correlation time-domain function, $k(\tau)$; the latter function models the exponential decay of the position of magnetic vectors with time due to Brownian motion, as defined by a statistical treatment of the loss of correlation by the ensemble of nuclear spins (Harris, 1983). In NMR, spin-lattice relaxation from the excited state to the ground state occurs when the magnetic moment of the nucleus interacts with magnetic fields fluctuating at its Larmor (or resonance) frequency. Molecular tumbling due to Brownian motion is at the origin of almost all relaxation mechanisms. Fluctuating magnetic fields originate from nuclear magnetic moments, unpaired electrons, electronic clouds, etc. Energy transfers resulting in relaxation will occur if the position vectors that determine instantaneous magnetic coupling of a nucleus with its surroundings are a function of time and the variations between these "position" functions over short intervals of time are characterized by the auto-correlation function, $k(\tau)$, which in turn, is characterized by the correlation time, τ_c. When the auto-correlation function is Fourier transformed, it is converted to a frequency domain function known as the spectral density function, $J(\omega)$, which describes the density of single quantum transitions possible at a particular frequency. The efficiency of longitudinal relaxation (T_1) is affected by the external magnetic field strength (Reisse, 1982; Harris, 1983). The relaxation is more favoured if $J(\omega)$ is large and the relaxation occurs for motions at the Larmor frequency, where $J(\omega_o)$ is proportional to $\tau_c/(1+\omega_o^2\tau_c^2)$ and also to the relaxation rate T_1^{-1}. T_1 is most efficient when $\tau_c = 1/\omega_o$.

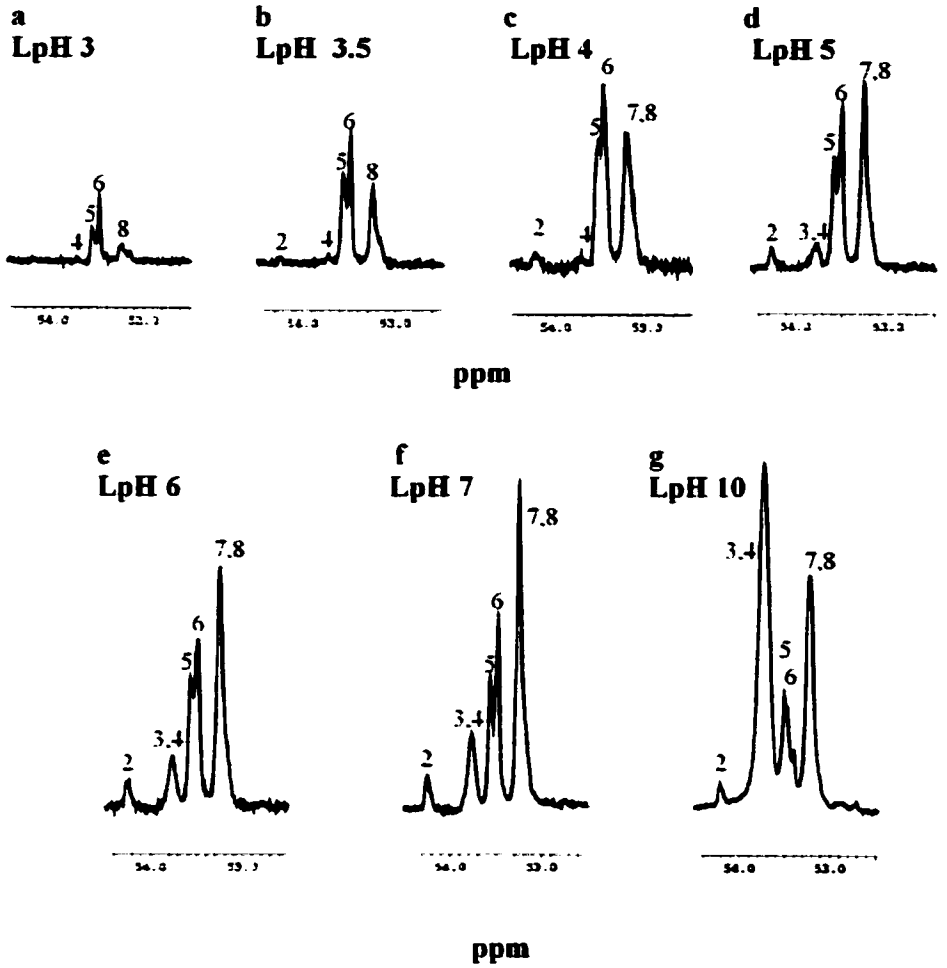


Figure 3-2. Expanded ^{13}C NMR spectra of ^{13}C-methylated ribonuclease A at various LpH values (LpH 3-10; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) unidentified small molecule (56.21 ppm); (2) unidentified α-$^{+}NMe_3$ impurity (54.23 ppm); (3) $Me_3N^{\varepsilon+}$-Lys (53.76 ppm); (4) C_αOOMe (53.71 ppm); (5) C_γOOMe (53.56 ppm); (6) C_γOOMe (53.48 ppm); (7) $Me_3N^{\alpha+}$-Lys (53.24 ppm); (8) C_δOOMe (53.23 ppm).

dialyzed. It was observed to decrease with LpH and with prolonged hydroxylamine treatment. The peak at 54.23 ppm appears to arise from a trimethylamino group, since it remains after hydroxylamine treatment. Both peaks are also present in the spectra of methylated α-chymotrypsin at 54.21 ppm and 56.22 ppm. These two peaks were also present in spectra of methylated soybean trypsin inhibitor (see Chapter 2). Neither peak was present in spectra of methylated human albumin, ovalbumin, lysozyme or insulin.

[It can be speculated that the resonance at 56.21 ppm is an $-O^{13}CH_3$, possibly due to methylation of the peptide backbone, namely the oxygen atom of the carbonyl group of peptide bonds (Blackburn, Hemphill Carter, and Phillips, 1941); however, the linewidth of the resonance was very narrow compared to resonances deduced to arise from derivatized protein functional groups. If the resonance at 56.21 ppm arises from the modification of the peptide backbone, it may be that the resonance disappears after dialysis in acidic water due to its instability rather than the loss of carbon-13 labeled small molecules. The fact that the resonance is smaller progressively with LpH may again be indicative of an unstable derivative at high pH. An imido group that undergoes salt formation (forming *O*-methylimidinium salt) by reaction with iodomethane would be expected to hydrolyze freely and to lose HI at pH 5 or greater.]

Figure 3-1A (Appendix 3-2) shows the expanded ^{13}C-NMR spectra of ^{13}C-methylated ribonuclease A (LpH 5), reacted *in vacuo* with ^{13}C-iodomethane at various times. Incorporation of ^{13}C-label into the protein was readily detected after 30 min and was essentially complete after 24 h.

Ribonuclease A has five Asp and five Glu residues in its polypeptide chain (Dayhoff, 1972). The intensity of the resonances of the ^{13}C-methyl esters of the γ- and

δ-carboxyl groups is much greater than the ester of the solitary α-carboxyl groups. At LpH values between 3-5 (Figure 3-2), the relative proportion of the ^{13}C-resonances of the γ- and δ-carboxyl ^{13}C-methyl esters formed by reaction with ^{13}C-iodomethane reflects the pK_a values of the respective carboxyl groups. At LpH 3, 3.5, and 4, the γ-carboxyl groups react more readily than the δ-carboxyl groups. At LpH 4.5 and 5, the relative proportions of reacted side-chain carboxyl groups of Asp and Glu residues balance, as both types of carboxyl groups are almost completely deprotonated and are present in the same amounts (i.e., five residues each). The α-carboxyl group (C-terminal Val; 53.71 ppm) reacts to a small extent with ^{13}C-iodomethane at LpH 3.5, and to a greater extent up to LpH 6-7. At higher LpH values the resonance intensities of the α- and ε-trimethylated amino groups increase. The resonance of the ^{13}C-trimethylated α-amino group of the N-terminal lysyl residue overlaps that of the δ-carboxyl ^{13}C-methyl esters; and the resonance of the ^{13}C-trimethylated ε-amino groups of lysyl residues (53.76 ppm) overlaps that of the methylated α-carboxyl group. The latter makes it difficult to make any qualitative assessment on the extent of reaction of the α-carboxyl methyl esters from the ^{13}C-NMR spectra. Quantitative results on the extent of methylation of carboxyl groups were obtained from difference titration data (see Table 3-5 below).

The ^{13}C-resonances of the side-chain carboxyl methyl esters of either Asp or Glu residues appear to be less influenced by their environment compared to those of the α-carboxyl methyl esters, as they overlap each other in their respective chemical shift ranges. For methylated ribonuclease A, this appears to be more or less true for the resonances of the ^{13}C-methyl esters of Glu residues. However, the environment of the side-chain carboxyl methyl esters of Asp residues appears to influence the chemical shifts

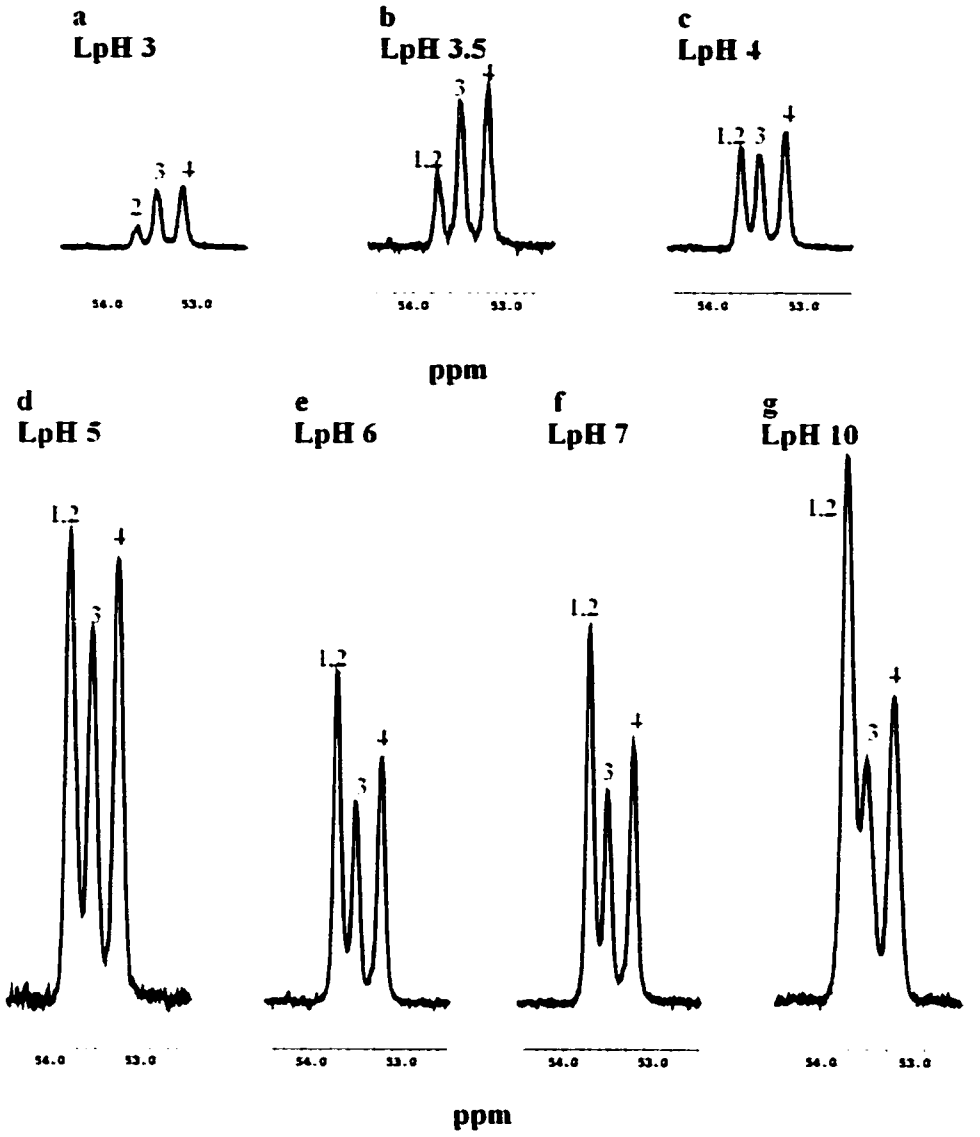


Figure 3-3. Expanded ^{13}C NMR spectra of ^{13}C-methylated human albumin at various LpH values (LpH 3-10; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) $\mathbf{Me_3}N^{\varepsilon+}$-Lys (53.76 ppm); (2) $C_\alpha OO\mathbf{Me}$ (53.69 ppm); (3) $C_\gamma OO\mathbf{Me}$ (53.49 ppm);

(4) $C_\delta OO\mathbf{Me}$ (53.19 ppm).

of these resonances. When the magnet of the NMR instrument was shimmed properly, it was possible to observe distinct, yet overlapping, peaks for the γ-carboxyl methyl esters of Asp residues at 53.56 and 53.48 ppm (Figure 3-2).

3.2.3 Human Albumin

Figure 3-3 shows the ^{13}C-NMR spectra of ^{13}C-methylated human albumin, obtained by reaction with ^{13}C-iodomethane at LpH values between 3 and 10. Figure 3-3 demonstrates the preferential methylation of the α-, γ-, and δ-carboxyl groups of human albumin reacted between LpH 3 and 5. The chemical shifts of the ^{13}C-resonances corresponding to the carboxyl methyl esters are: 53.65 (α), 53.49 (γ), and 53.19 ppm (δ).

Human albumin has 35 Asp and 62 Glu residues in its polypeptide chain (Dayhoff, 1972). Between LpH 3 and 5 (Figure 3-3), the relative proportion of the ^{13}C-resonances of the γ- and δ-carboxyl ^{13}C-methyl esters formed by reaction with ^{13}C-iodomethane reflects the pK_a values and the amounts of the respective carboxyl groups. At LpH 3 and 3.5, the γ-carboxyl groups should react more readily than the δ-carboxyl groups. This is not obvious from the ^{13}C-NMR sprectra, since the intensity of the resonances for these groups appear to be the same. The ratio of the number of Glu to Asp residues is approximately 1.8:1. If the ratio were 1:1, as in the case of ribonuclease A, the intensity of the resonances for the γ-carboxyl ^{13}C-methyl esters would be approximately twice that for the δ-carboxyl ^{13}C-methyl esters. However, the similar resonance intensities indicate that there are more δ- than γ-carboxyl ^{13}C-methyl esters. Both types of carboxyl groups are almost completely ionized between LpH 4.5 and 5. At higher LpH values (≥ LpH 4), it is observed that the relative proportion of the

resonances of the δ-carboxyl ^{13}C-methyl esters is greater than that of the γ-carboxyl ^{13}C-methyl esters, reflecting their relative amounts.

The α-carboxyl group [C-terminus Leu-609 (53.69 ppm)] reacts to a small extent with ^{13}C-iodomethane at LpH 3.5, but to a greater extent at higher LpH values (≥LpH 4). However, the ^{13}C-resonances from the 57 N^{ϵ}-trimethylated Lys residues (53.76 ppm) begin to overlap those from the α-carboxyl methyl esters, thus making it difficult to make any qualitative assessment on the extent of reaction of the latter from the ^{13}C-NMR spectra. Interestingly, at LpH values between 3 and 3.5 the broad resonance due to several overlapping N^{ϵ}-trimethyl amino groups of lysyl residues suggests that a significant proportion of nucleophilic unprotonated amino groups is present in the lyophilized protein structure. An alternative explanation is that protonated amino groups may react in the *in vacuo* environment by first transferring their proton to a counterion, followed by nucleophilic attack of iodomethane (see discussion).

3.2.4 α-Chymotrypsin

Figure 3-2A (Appendix 3-2) shows the expanded ^{13}C-NMR spectra of α-chymotrypsin (LpH 5) reacted *in vacuo* (75°C) at various times. The time course experiment shows that incorporation of ^{13}C-label into the protein was readily detected after 30 min, and was essentially complete after 24 h. Figure 3-4 shows the ^{13}C-NMR spectra of ^{13}C-methylated α-chymotrypsin, obtained by reaction with ^{13}C-iodomethane at LpH values between 3 and 10. Figure 3-4 demonstrates the preferential methylation of the α-, γ-, and δ-carboxyl groups of α-chymotrypsin reacted between LpH 3 and 5. The chemical shifts of the ^{13}C-resonances corresponding to the carboxyl methyl esters are:

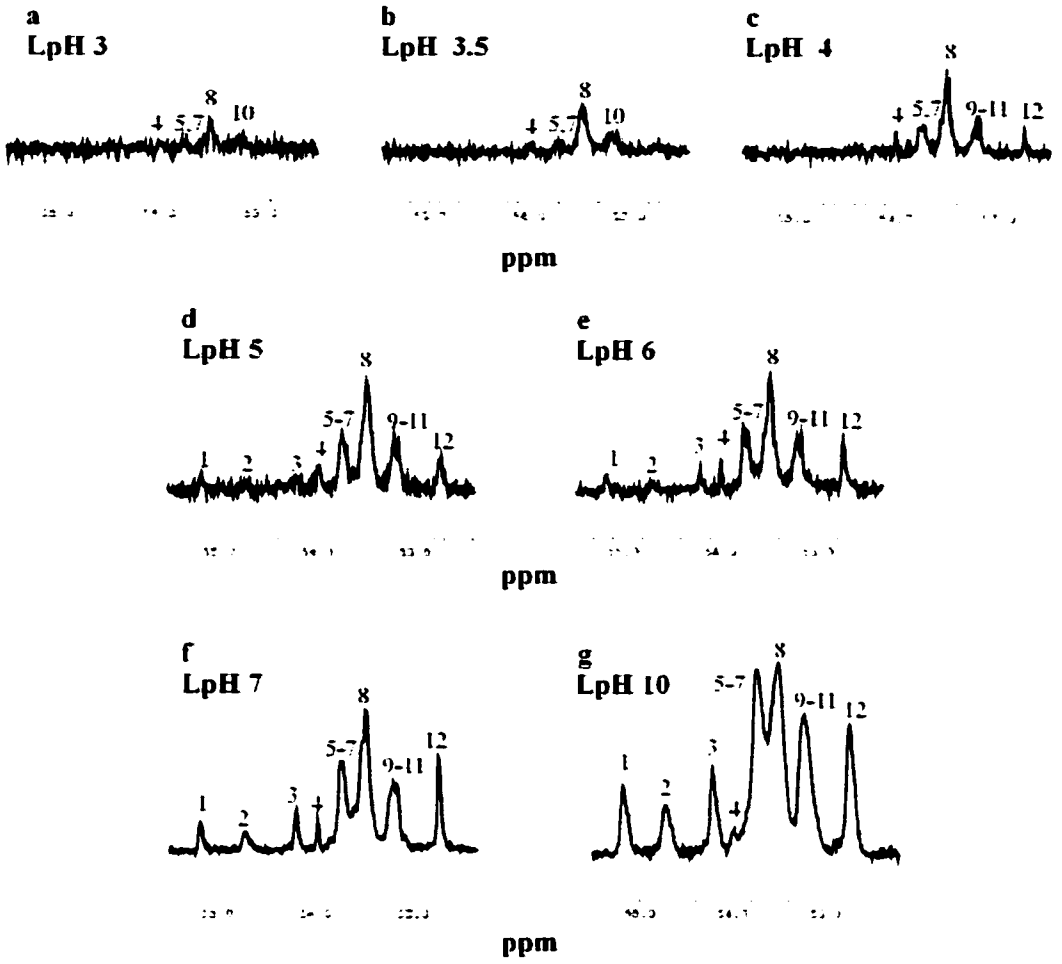


Figure 3-4. Expanded ^{13}C NMR spectra of ^{13}C-methylated α-chymotrypsin at various LpH values (LpH 3-10; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1), (2), (3) unidentified α-$^{+}NMe_3$, impurities (55.16 ppm, 54.72 ppm, and 54.19 ppm);

(4) $C_\alpha OOMe$ (53.99 ppm; C-terminal Asn-245); (5) $C_\alpha OOMe$ (53.77 ppm);

(6) $Me_3N^{\varepsilon+}$-Lys (53.74 ppm); (7) $C_\alpha OOMe$ (53.72 ppm); (5, 7) C-termini: Leu-13, Tyr-146; (8) $C_\gamma OOMe$ (53.50 ppm); (9) half-cystine-$N^{\alpha+}Me_3$; (10) $C_\delta OOMe$ (53.21 ppm); (11) $Me_3N^{\alpha+}$-Ile (53.17 ppm); (12) $Me_3N^{\alpha+}$-Ala (52.74 ppm).

53.99 (α), 53.77 (α), 53.72 (α), 53.50 (γ), and 53.21 ppm (δ). There may be more than two resonances due to α-carboxyl methyl esters between 53.72-53.77 ppm, since the peaks overlap and are broadened. Other evidence for the latter is that at higher LpH values, there are three small peaks that correspond to N^{α}-trimethylated amino acid residues (55.16 ppm, 54.72 ppm, and 54.19 ppm), which are most likely impurities due to autoproteolysis, and which do not disappear from the spectrum with hydroxylamine treatment of the sample.

The α-amino groups also react to a small extent to form the trimethylated amino acid residues at LpH values less than 4). The chemical shifts of the ^{13}C-resonances corresponding to these derivatives are: 53.4 ppm for half-cystine (overlapped by the γ-carboxyl methyl ester resonances at 53.50 ppm), 53.17 ppm for N^{α}-trimethylated Ile, and 52.74 ppm for N^{α}-trimethylated Ala.

α-Chymotrypsin has nine Asp and five Glu residues in its three polypeptide chains (Dayhoff, 1972). Between LpH 3 and 5 (Figure 3-4), the relative proportion of the ^{13}C-resonances of the γ- and δ-carboxyl methyl esters formed by reaction with ^{13}C-iodomethane reflects the pK_a values of the respective carboxyl groups. At LpH 3, 3.5, and 4, the γ-carboxyl groups react more readily than the δ-carboxyl groups. At LpH 4.5 and 5, the relative proportions of reacted side-chain carboxyl groups of Asp and Glu residues do not balance, as in the case of ribonuclease A, but rather appear to reach their optimal relative reactivities (i.e., the resonances for the overlapped Asp carboxyl methyl esters are larger than those for the overlapped Glu carboxyl methyl esters). Both types of carboxyl groups are almost completely ionized between LpH 4.5 and 5 and the

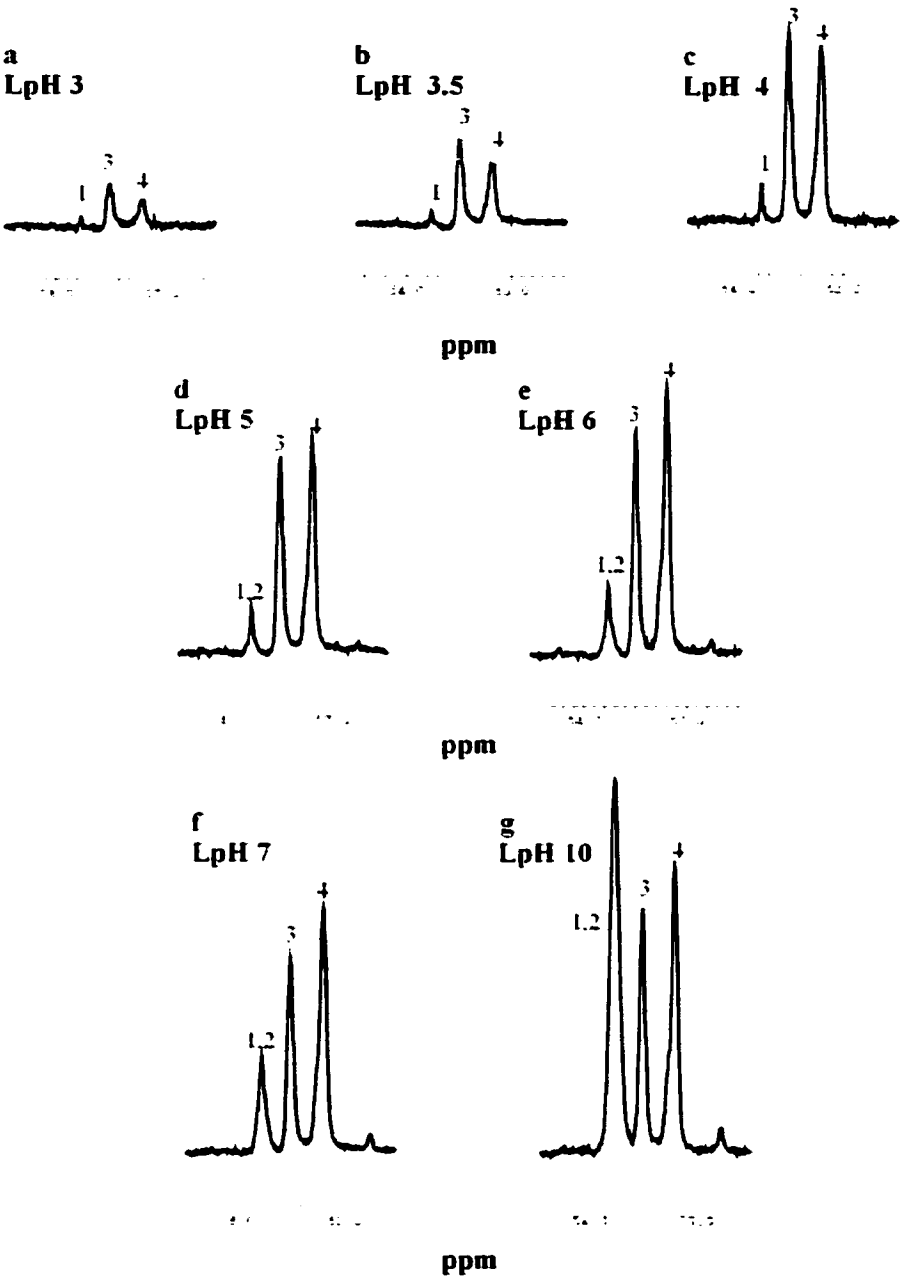


Figure 3-5. Expanded ^{13}C NMR spectra of ^{13}C-methylated ovalbumin at various LpH values (LpH 3-10; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) C_{α}OO**Me** (53.77 ppm. C-terminal Pro-385); (2) **Me**$_3$N$^{\varepsilon+}$-Lys (53.75 ppm);

(3) C_{γ}OO**Me** (53.49 ppm); (4) C_{δ}OO**Me** (53.19 ppm).

relative proportion of each resonance reflects the amount of each amino acid residue. The α-carboxyl groups [C-termini: Leu-13, Tyr-146 (53.72-53.77 ppm), Asn-245 (53.99 ppm)] react to a small extent with ^{13}C-iodomethane at LpH 3.5, but to a greater extent at higher LpH values (LpH 6-7). However, the ^{13}C-resonances from N^{ε}-trimethylated Lys residues (53.74 ppm) begin to overlap those from the α-carboxyl methyl esters, thus making it difficult to make any qualitative assessment on the extent of reaction of the latter from the ^{13}C-NMR spectra.

3.2.5 Ovalbumin

Figure 3-5 shows the ^{13}C-NMR spectra of ^{13}C-methylated ovalbumin, obtained by reaction with ^{13}C-iodomethane at LpH values between 3 and 10. The results demonstrate the preferential methylation of the α-, γ-, and δ-carboxyl groups of ovalbumin reacted between LpH 3 and 5. The chemical shifts of the ^{13}C-resonances corresponding to the carboxyl methyl esters are: 53.76 (α), 53.49 (γ), and 53.19 ppm (δ).

Ovalbumin has 14 Asp and 33 Glu residues in its polypeptide chain (Dayhoff, 1972). Between LpH 3 and 5 (Figure 3-5), the relative proportion of the ^{13}C-resonances of the γ- and δ-carboxyl ^{13}C-methyl esters formed by reaction with ^{13}C-iodomethane reflects the pK_a values and the amounts of the respective carboxyl groups. At LpH 3 and 3.5, the γ-carboxyl groups should and do react more readily than the δ-carboxyl groups. At LpH 4.5, the intensity of the resonances for these groups is the same, as expected since the ratio of the number of Glu to Asp residues is approximately 2.4:1. If the ratio were 1:1, as in the case of ribonuclease A, the intensity of the resonances for the γ-carboxyl ^{13}C-methyl esters would be approximately twice that for the δ-carboxyl ^{13}C-methyl esters. However, the similar resonance intensities indicate that there are more

δ- than γ-carboxyl ^{13}C-methyl esters. Both types of carboxyl groups are almost completely ionized between LpH 4.5 and 5. At higher LpH values (≥ 5), it is observed that the relative proportion of the resonances of the δ-carboxyl ^{13}C-methyl esters is greater than that of the γ-carboxyl ^{13}C-methyl esters, reflecting their relative amounts.

The α-carboxyl group [C-terminus Pro-385 (53.76 ppm)] reacts to a small extent with ^{13}C-iodomethane at LpH 3 and 3.5, but to a greater extent at higher LpH values (≥LpH 4). However, the ^{13}C-resonances from the 18 N^{ε}-trimethylated Lys residues (53.75 ppm) begin to overlap those from the α-carboxyl methyl esters, thus making it difficult to make any qualitative assessment on the extent of reaction of the latter from the ^{13}C-NMR spectra. There was no evidence of a resonance due to a ^{13}C-trimethylated amino group, which reflected the fact that ovalbumin has an acetylated N-terminal Gly residue. As with human albumin, for ovalbumin at LpH values between 3 and 3.5, the broad resonance due to several overlapping N^{ε}-trimethyl amino groups of lysyl residues suggests that a significant proportion of nucleophilic unprotonated amino groups is present in the lyophilized protein structure. It is possible that protonated amino groups may react by first transferring their proton to a counterion, followed by nucleophilic attack of iodomethane (see discussion).

3.2.6 Lysozyme

Figure 3-6 shows the ^{13}C-NMR spectra of ^{13}C-methylated chicken egg-white lysozyme, obtained by reaction with ^{13}C-iodomethane at LpH values between 3 and 10. The spectra demonstrate the preferential methylation of the α-, γ-, and δ-carboxyl groups of lysozyme reacted between LpH 3 and 5. The chemical shifts of the ^{13}C-resonances

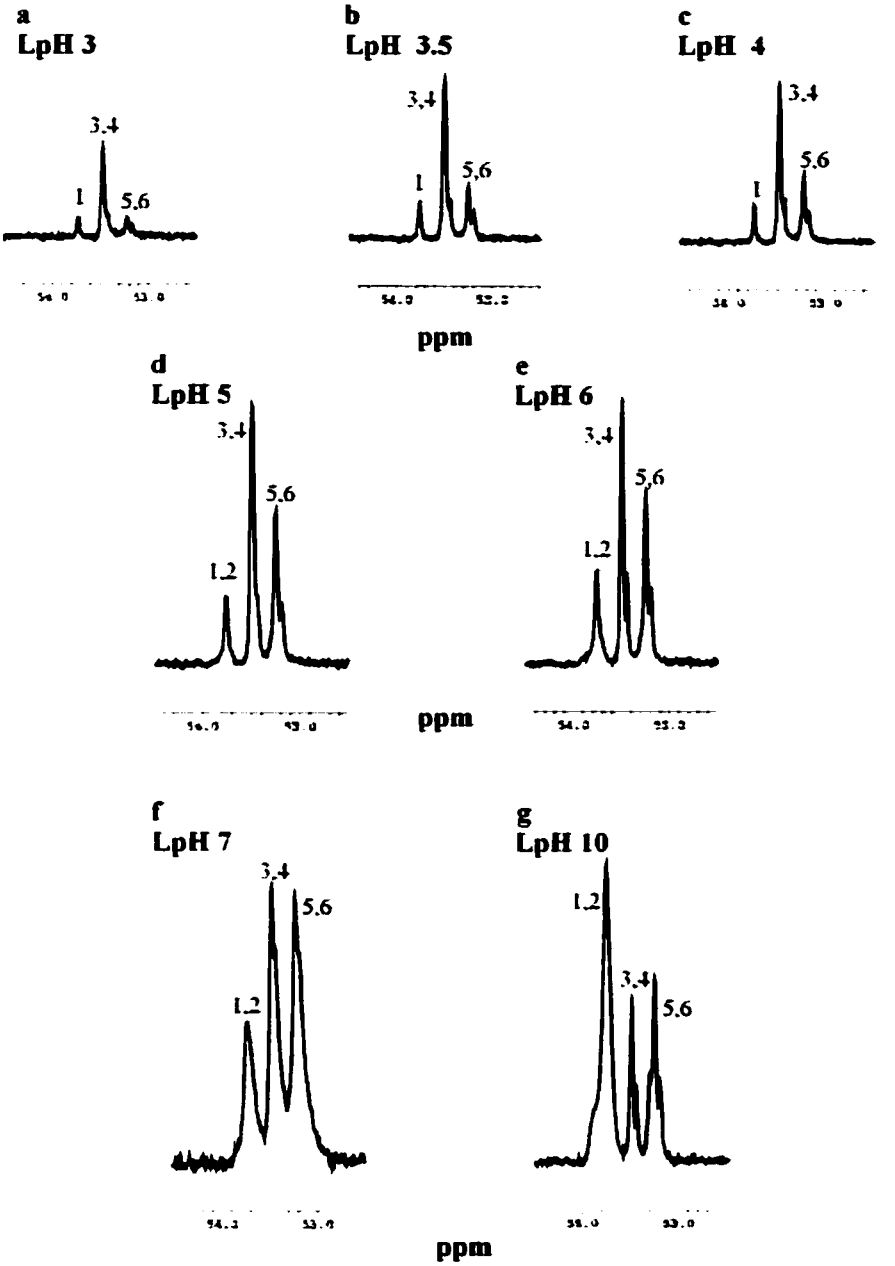


Figure 3-6. Expanded ^{13}C NMR spectra of ^{13}C-methylated chicken egg-white lysozyme at various LpH values (LpH 3-10; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) C_α**OOMe** (53.75 ppm, C-terminal Leu-147); (2) **Me**$_3$**N**$^{\varepsilon+}$-Lys (53.75 ppm); (3) C_γ**OOMe** (53.47 ppm); (4) C_γ**OOMe** (53.41 ppm); (5) C_δ**OOMe** (53.22 ppm); (6) C_δ**OOMe** (53.16 ppm).

corresponding to the carboxyl methyl esters are: 53.73 (α), 53.47 (γ), 53.41 (γ), 53.22 (δ), and 53.16 ppm (δ).

Lysozyme has seven Asp and two Glu residues in its polypeptide chain (Dayhoff, 1972). Figure 3-6 shows that between LpH 3 and 5 the relative proportion of the ^{13}C-resonances of the γ- and δ-carboxyl ^{13}C-methyl esters formed by reaction with ^{13}C-iodomethane reflects the pK_a values (and the relative amounts) of the respective carboxyl groups. At LpH 3, 3.5, and 4, the γ-carboxyl groups react more readily than the δ-carboxyl groups. At LpH 4.5 and higher, the relative proportions of reacted side-chain carboxyl groups of Asp and Glu residues appear to be disproportionate with respect to their relative amounts. However, this is because the resonance of the N^{α}-trimethylamino group of the N-terminal Lys residue begins to overlap that of the δ-carboxyl ^{13}C-methyl esters (Glu). Both types of carboxyl groups are almost completely ionized at LpH 5 and are present in the ratio of 3.5:1 Asp:Glu. The α-carboxyl group (C-terminal Leu-147; 53.75 ppm) reacts to a small extent with ^{13}C-iodomethane at LpH 4.5, and to a greater extent up to LpH 6-7. However, the ^{13}C-resonances from N^{ϵ}-trimethylated Lys residues (53.75 ppm) begin to overlap those from the α-carboxyl ^{13}C-methyl esters, thus making it difficult to make any qualitative assessment on the extent of reaction of the latter from the ^{13}C-NMR spectra.

In general, the ^{13}C-resonances of the side-chain carboxyl methyl esters of either Asp or Glu residues appear to be less influenced by their environment compared to those of the α-carboxyl methyl esters, as they overlap each other in their respective chemical shift ranges. For methylated lysozyme, this appears to be more or less true at higher LpH values (≥6.5); however, at lower LpH values the environment of the side-chain

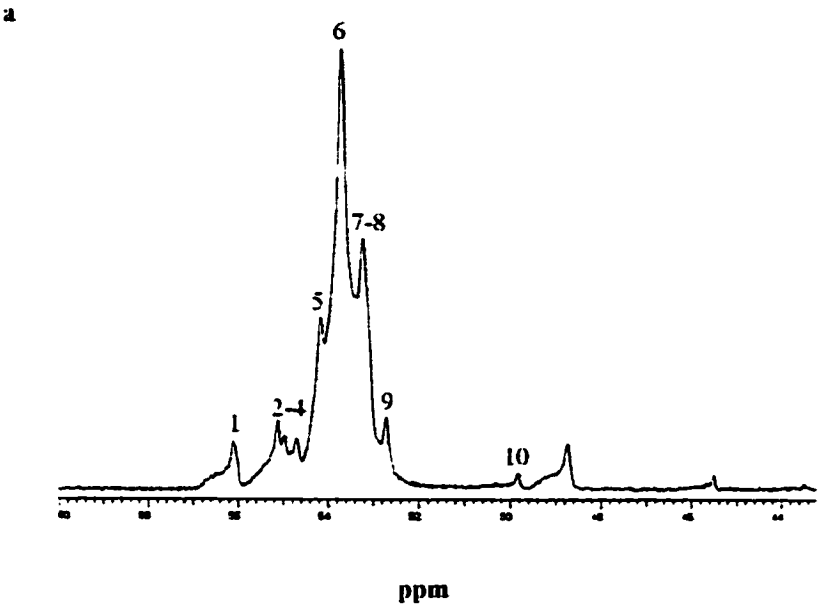


Figure 3-7(a). Expanded ^{13}C NMR spectra (14208 scans) of ^{13}C-methylated *Bacillus thuringiensis* insecticidal toxin (T_1, T_2; *Bt*) (LpH 10; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) Tyr-O**Me**; (2), (3), (4) $\mathbf{Me_3N^{\alpha+}}$-impurities; (5) C_αO**OMe**;

$\mathbf{Me_3N^{\alpha+}}$-impurity; (6) $\mathbf{Me_3N^{\varepsilon+}}$Lys; (7), (8) $C_{\gamma,\delta}$O**OMe**; (9) $\mathbf{Me_3N^{\alpha+}}$Ile;

(10) MeOH.

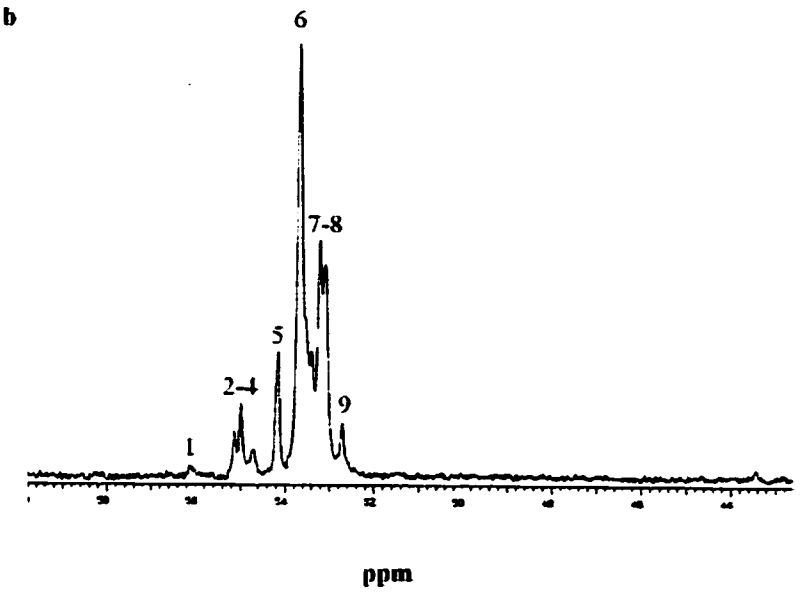


Figure 3-7(b). Expanded ^{13}C NMR spectra (14208 scans) of ^{13}C-methylated *Bacillus thuringiensis* insecticidal toxin (T_1 and T_2; *Bt*) after (b) dialysis against 8M urea (followed by dialysis against water and prepared for NMR analysis).

The spectrum of dialyzed ^{13}C-methylated *Bt* toxin illustrates the presence of α-, γ-, and δ-carboxyl ^{13}C-methyl esters (53.6-54.2, 53.4 and 53.2 ppm, respectively), as they are easily hydrolyzed (albeit incompletely) in the slightly alkaline medium of urea. Peak resonances correspond to the ^{13}C-enriched methyl groups in: (1) Tyr-O**Me**; (2), (3), (4) $\mathbf{Me_3N^{\alpha+}}$-impurities; (5) C_αOO**Me**; $\mathbf{Me_3N^{\alpha+}}$-impurity; (6) $\mathbf{Me_3N^{\varepsilon+}}$Lys; (7), (8) $C_{\gamma,\delta}$OO**Me**; (9) $\mathbf{Me_3N^{\alpha+}}$Ile; (10) MeOH.

carboxyl methyl esters of Asp or Glu residues appears to influence the chemical shifts of these resonances. When the magnet of the NMR instrument was shimmed properly, it was possible to observe distinct, yet overlapping, peaks for both the γ- and δ-carboxyl methyl esters (Figure 3-6).

3.2.7 *Bt* toxin (*Bacillus thuringiensis*)

Figure 3-7 shows the ^{13}C-NMR spectra of ^{13}C-methylated *Bacillus thuringiensis* toxin (T_1 and T_2; *Bt*, LpH 10) before and after dialysis against 8 M urea. Although the discovery that carboxyl groups of protein react with iodomethane *in vacuo* had not been observed at the time of these analyses, the spectrum of dialyzed ^{13}C-methylated *Bt* toxin illustrates the presence of α-, γ-, and δ-carboxyl ^{13}C-methyl esters (53.6-54.2, 53.4 and 53.2 ppm, respectively), as they are easily hydrolyzed (albeit incompletely) in the slightly alkaline medium of urea.

Table 3-3 summarizes the chemical shifts of α-, γ-, and δ-carboxyl ^{13}C-methyl ester resonances in the various *in vacuo* ^{13}C-methylated proteins studied. The data clearly illustrates that the resonances for α-, γ- and δ-carboxyl ^{13}C-methyl esters reside in specific and distinct regions of the ^{13}C-NMR spectrum.

3.2.8 Explanation of the Differences in Chemical Shifts of the α-, γ-, and δ-Carboxyl Methyl Esters in ^{13}C-Methylated Proteins

The example presented below using ethylenediamine is put forth in order to explain the differences in pK_a values of α-, γ-, and δ-carboxyl groups in proteins. The electron-withdrawing inductive effect of the peptide bond on the α-carboxyl methyl ester functionality can explain why the C-terminal carboxyl methyl ester ^{13}C-resonances are more downfield and distinct compared to those of the side-chain carboxyl methyl esters of Asp and Glu residues. The α-carboxyl methyl ester resonances are influenced by the

Table 3-3 Chemical Shifts of α-, γ-, δ-Carboxyl ^{13}C-Methyl Ester Resonances in *in vacuo* ^{13}C-Methylated Proteins

^{13}C-methylated protein	Chemical shifts of $C_\alpha OO^{13}Me$ (ppm)	Chemical shifts of $C_\gamma OO^{13}Me$ (ppm)	Chemical shifts of $C_\delta OO^{13}Me$ (ppm)
bovine insulin	A-chain Asn, 54.00; B-chain Ala, **53.76**		four overlapping Glu residues, 53.15
high purity insulin human injection sample	A-chain Asn, 54.00; B-chain Ala, **53.75**		four overlapping Glu residues, 53.16
recombinant mutant [Lys(B28), Pro(B29)] insulin injection sample	A-chain Asn, 54.00; B-chain Ala, **53.83**		four overlapping Glu residues, 53.14
ribonuclease A	53.71	53.56, 53.48	53.23
human albumin	53.69	53.49	53.19
α-chymotrypsin	C-terminal Asn-245, 53.99; C-termini: Leu-13, Tyr-146, 53.77, 53.72	53.50	53.21
ovalbumin	C-terminal Pro-385, 53.77	53.49	53.19
lysozyme	C-terminal Leu-147, 53.75	53.47, 53.41	53.22, 53.16
***Bt* toxin**	53.6-54.2	53.4	53.2

functionality of the side-chain (see resonances for dipeptide standards, Table 3-2). The ^{13}C-resonances of the side-chain carboxyl methyl esters of Asp residues are affected more by the polar character of the peptide bond compared to those of the Glu residues. This is evidenced by the former resonances residing more downfield compared to those of the Glu residues. Apart from the latter, the ^{13}C-resonances of the side-chain carboxyl methyl esters of either Asp or Glu residues appear to be less influenced by their environment, as the individual resonances overlap each other in their respective chemical shift ranges.

Ethylenediamine, $^{+}NH_3CH_2CH_2NH_3^{+}$, has a pK_{a1} of approximately 6 and a pK_{a2} of approximately 9 (Harris, 1987b). Why? The first ionization (which normally would be expected above pH 9-10) is encouraged (at least one thousand fold) by the unfavourable electrostatic repulsion of the neighbouring positive charge. The second ionization is much more like a normal pK_a value of an aliphatic primary amine because there is very little inductive or charge participation by the neighbouring (now neutral) amine. There is little influence on the accustomed ionization value.

The acids of Glu and Asp are so far apart from the rest of the protein that they more or less resemble acetic acid or an aliphatic primary carboxylic acid. Interestingly, the pK_a values of the side-chain carboxyl group of Glu and the carboxyl group of acetic acid are similar (4.42 and 4.76, respectively), while that for the side-chain carboxyl of Asp is 3.90 (in the free amino acid; Harris, 1987a). In the protein structure, the pK_a values of Asp and Glu side-chain carboxyl groups are typically in the ranges of 3.9-4.0 and 4.3-4.5, respectively (Creighton, 1993b). In contrast, the ionization constants of C-terminal acids are much more easily (2-10 times) ionized, as shown by the lower pK_a values in the range of 3.5-4.3 (Creighton, 1993b). This is because the ease of ionization

is influenced by the electropositivity of the carbonyl carbon. This carbon is directly bonded to the acidic oxygen that is developing a negative charge during the dissociation process. In the case of the C-termini, the peptide bond between the penultimate amino acid residue and that of the C-terminal residue has an inductive electron withdrawing effect on the C-terminal carbonyl carbon, owing to the partial double bond character of the peptide bond, which is presumably due to resonance (Creighton, 1993a). Electrons tend to redistribute about the peptide bond resulting in a permanent dipole moment in the peptide backbone. This makes the carbonyl carbon of the α-carboxyl group more electropositive than it normally would be from the effect of the oxygen-carbon double bond alone. The more electropositive carbon makes the acidic oxygen more likely to develop a negative charge to cancel the electropositive center, and the acid dissociation proceeds more readily with α-carboxyl groups.

3.3 *In Vacuo* Esterification of Model Proteins with ^{13}C-Methanol or ^{13}C-Ethanol and Hydrochloric Acid Catalyst

The *in vacuo* reaction of model proteins with ^{13}C-methanol or ^{13}C-ethanol and hydrochloric acid catalyst (4N HCl in dioxane) resulted in the exclusive formation of ^{13}C-methyl or ^{13}C-ethyl esters of carboxylic acid groups in protein, as evidenced by ^{13}C-NMR spectroscopy of the observed chemical shifts for α-, γ-, and δ-carboxyl ^{13}C-methyl or ^{13}C-ethyl ester resonances. The ^{13}C-methyl ester peaks were assigned from acetylated dipeptide ^{13}C-methyl ester standards, and assignments were comparable to those made for several proteins methylated with ^{13}C-iodomethane. When the modified protein was treated with base or hydroxylamine in aqueous solution, all the resonances disappeared with the concomitant appearance of the resonance for ^{13}C-methanol at 49.8 ppm, showing that they arose from esterification of the various classes of carboxyl

groups in the protein. Hydroxylamine reacted *in vacuo* for 4 h with ^{13}C-methyl-esterified ribonuclease A removed all ester resonances from the ^{13}C-NMR spectrum in the range of 52-57 ppm (Figure 3-8c). Peaks observed downfield of the methyl ester resonance region at 56.2 and 55.5 ppm are derivatives of small molecules present in the protein preparation and were removed from the spectrum by dialysis (Figure 3-8d).

As in the case of the model dipeptides, the chemical shifts of the resonances among the esterified α-carboxyl groups in the proteins studied showed a much greater variation than among those of the esterified γ- or δ-carboxyl groups, indicating that the chemical shift of the α-carboxyl ^{13}C-methyl ester resonance in proteins is much more sensitive to its environment than that for either of the γ- or δ-carboxyl ^{13}C-methyl esters.

Although acetylated ^{13}C-ethyl-esterified dipeptide standard compounds were not prepared to verify the assignment of the chemical shifts of ^{13}C-ethyl-esterified proteins, it is reasonable to assume that the peak assignments are correct based on those of acetylated ^{13}C-methyl-esterified dipeptide standards and ^{13}C-methyl-esterified proteins. The pattern of derivatization is similar, as would be expected if an alcohol of one additional methylene group (i.e., ethyl vs methyl) were used as reagent.

Ribonuclease A was fully soluble after esterification, while α-chymotrypsin and insulin required urea to effect complete solubilization.

3.3.1 Ribonuclease A

Figure 3-3A (Appendix 3-2) shows the expanded ^{13}C-NMR spectra of a typical time course experiment for the *in vacuo* esterification of protein carboxyl groups. In this case, ^{13}C-methyl-esterified ribonuclease A (LpH 3) was reacted *in vacuo* with ^{13}C-methanol/HCl at 75°C. Incorporation of ^{13}C-label into the protein was readily

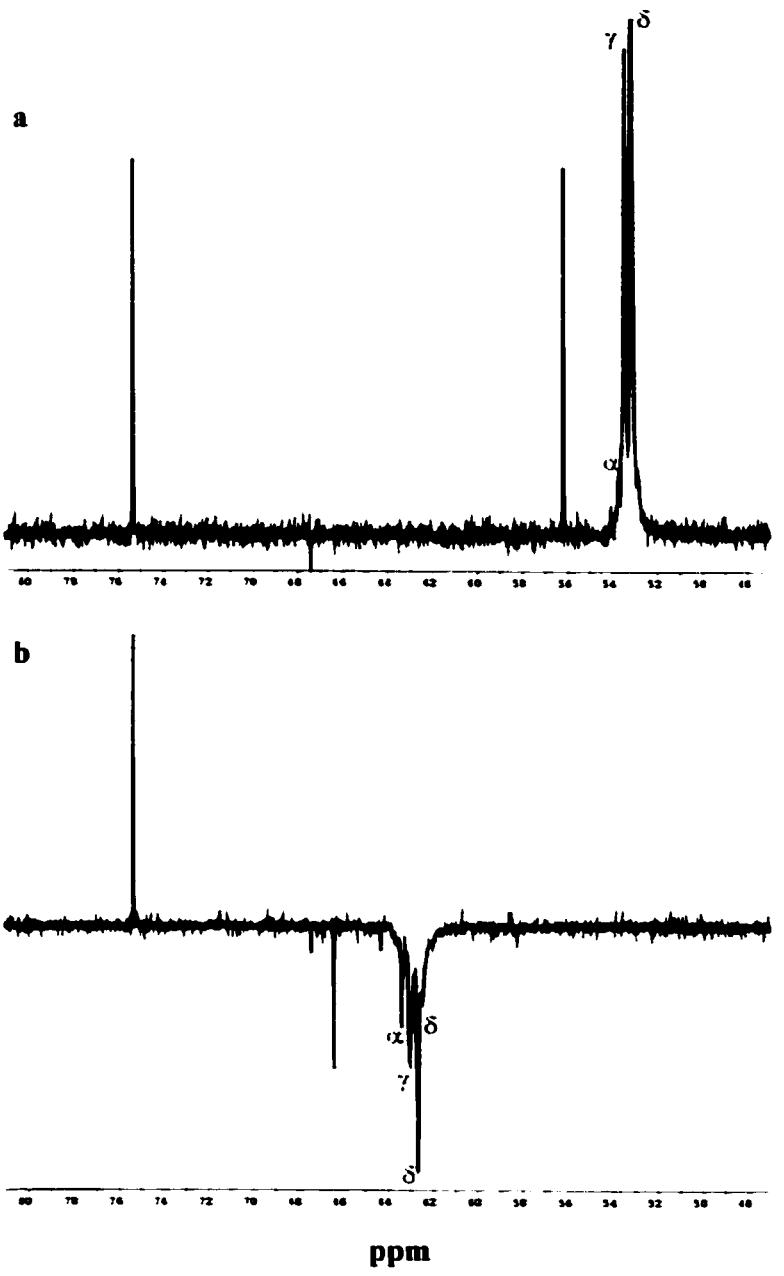


Figure 3-8. Expanded ^{13}C NMR spectra of (a) ^{13}C-methyl- and (b) ^{13}C-ethyl-esterified ribonuclease A at LpH 3 (reacted *in vacuo* with ^{13}C-methanol or ^{13}C-ethanol and hydrogen chloride catalyst).

Continued...

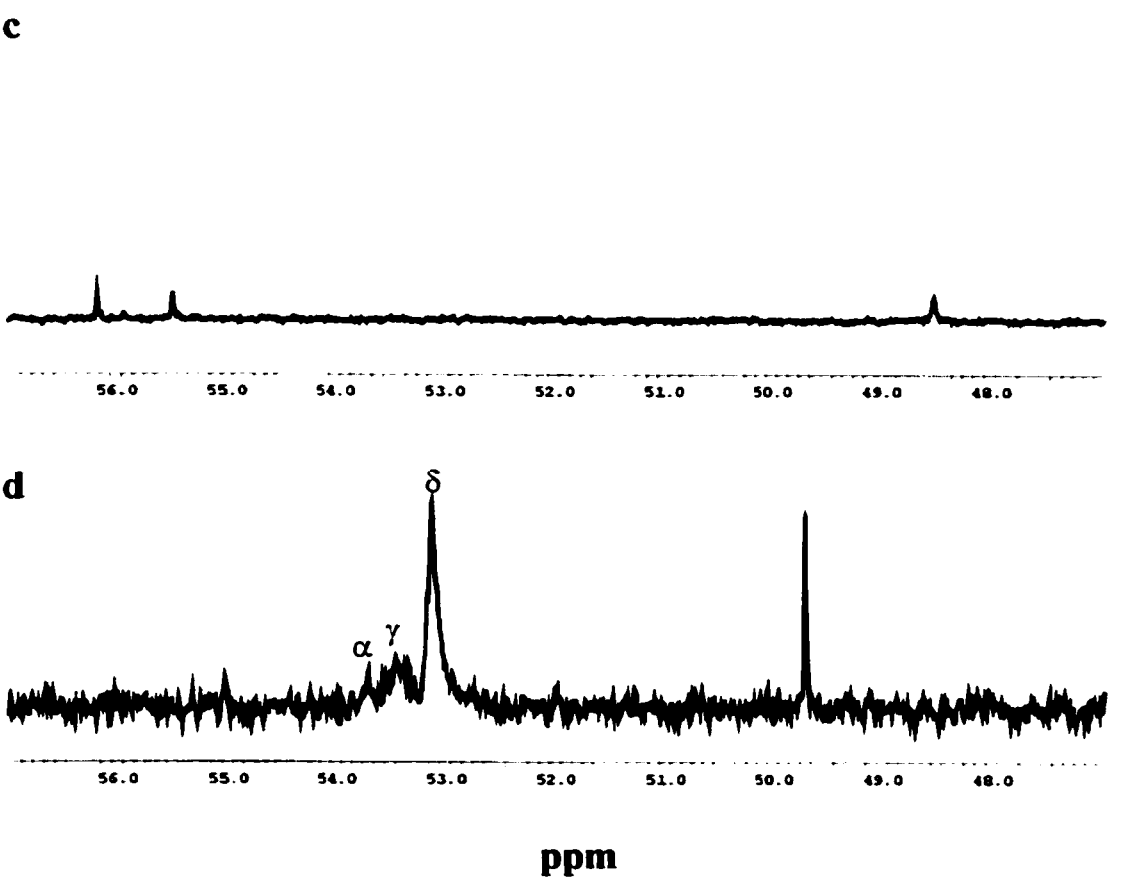


Figure 3-8. The selective nature of the esterification reaction for protein carboxyl groups is illustrated in (c), as the resonances for the carboxyl ^{13}C-methyl esters of ^{13}C-methyl-esterified ribonuclease A are removed by hydroxylamine reacted *in vacuo* for 4 h. Peaks observed downfield of the methyl ester resonance region at 56.2 and 55.5 ppm are derivatives of small molecules present in the protein preparation and are easily removed from the spectrum by dialysis (d).

The ^{13}C-resonances corresponding to the α-, γ-, and δ-carboxyl methyl esters in (a) are: 53.74 (α), 53.46 (γ), 53.41 (γ), and 53.18 ppm (δ) and those corresponding to the α-, γ-, and δ-carboxyl ethyl esters in (b) are: 63.37 (α), 63.02 (γ), 62.79 (γ), 62.64 (δ), and 62.45 ppm (δ). The resonances are negative in (b) because the ^{13}C-ethanol is enriched with carbon-13 at the C-1 position, which is a methylene group (-$^{13}CH_2$-), and show up as negative peaks in a DEPT-135 experiment. As observed with ethyl-esterified insulin (see Figure 3-9b), the resonance at 62.79 ppm may correspond to the γ-carboxyl ^{13}C-ethyl ester of an Asp residue formed from an Asn residue by methanolysis of its side chain amide group (Means and Feeney, 1971).

detected after 15 min and was essentially complete after 4 h, based on integrated peak areas. Integrated peak areas for the various times (15, 30 min, 1, 2, 4, 8, 12, 24, 48, and 72 h) were as follows, respectively: 149, 287, 572, 822, 968, 922, 1016, 1228, 1029, and 936. The initial rate of incorporation of ^{13}C-label is linear. At longer times, there is an increase in intensity of a sharp peak at 53.6 ppm. This resonance likely results from the methylation of a small molecule, as it disappears with dialysis, leaving the resonance for the single α-carboxyl ^{13}C-methyl ester, and those for the γ- and δ-carboxyl ^{13}C-methyl esters (Figure 3-8(d)).

Figure 3-8a shows the ^{13}C-NMR spectrum of *in vacuo* ^{13}C-methyl-esterified ribonuclease A at LpH 3, and clearly illustrates the three types of ^{13}C-methyl esters of protein α-, γ-, and δ-carboxyl groups. The chemical shifts of the ^{13}C-resonances corresponding to the carboxyl methyl esters are: 53.74 (α), 53.46 (γ), 53.41 (γ), and 53.18 ppm (δ). It is important to note that one resonance was observed in the α-carboxyl ^{13}C-methyl ester region of the spectrum. It will be seen later with the results from consecutive reactions (or longer reaction times; see Figure 3-3A, Appendix 3-2) that several small resonances appear in the latter region of the spectrum, which indicates that the peptide backbone may be cleaved under the acidic reaction conditions, leading to the esterification of the newly formed carboxylic acid groups (see section on extent of esterification of ribonuclease A below; also see description on esterified insulin below).

Figure 3-8b shows the ^{13}C-NMR spectrum of *in vacuo* ^{13}C-ethyl-esterified ribonuclease A at LpH 3. The spectrum illustrates the three types of ^{13}C-ethyl esters of protein α-, γ-, and δ-carboxyl groups; the resonances are negative because the ^{13}C-ethanol is enriched with carbon-13 at the C-1 position, which is a methylene group (-$^{13}CH_2$-).

These groups show up as negative peaks in a DEPT-135 experiment. The chemical shifts of the ^{13}C-resonances corresponding to the carboxyl ethyl esters are: 63.37 (α), 63.02 (γ), 62.79 (γ), 62.64 (δ), and 62.45 ppm (δ).

The ^{13}C-resonances of the side-chain carboxyl methyl esters of either Asp or Glu residues appear to be less influenced by their environment compared to those of the side-chain carboxyl ethyl esters of the same groups. The former resonances overlap each other in a single broad peak, while the latter resonances are observed as more resolved, but still overlap each other and may be due to the bulkier ethyl group. Ribonuclease A has five Asp and five Glu residues but only a single relatively broad resonance is observed for each of the corresponding γ- and δ-carboxyl ^{13}C-methyl esters of these amino acid side-chains (Figure 3-8). Similarly, for the esterification reaction with ^{13}C-ethanol/HCl, a single broad resonance is observed for the γ-carboxyl ^{13}C-ethyl esters of Asp residues. Two distinct resonances are observed for the δ-carboxyl ^{13}C-ethyl esters of the glutamate residues differing by approximately 0.05 ppm.

As observed with ethyl-esterified insulin (see Figure 3-9b), the resonance at 62.79 ppm may correspond to the γ-carboxyl ^{13}C-ethyl ester of an Asp residue formed from an Asn residue by methanolysis of its side-chain amide group (Means and Feeney, 1971).

3.3.2 Insulin

Figure 3-9a shows the ^{13}C-NMR spectrum of *in vacuo* ^{13}C-methyl-esterified insulin at LpH 3, and clearly illustrates the three types of ^{13}C-methyl esters of protein α-, γ-, and δ-carboxyl groups. The chemical shifts of the ^{13}C-resonances corresponding

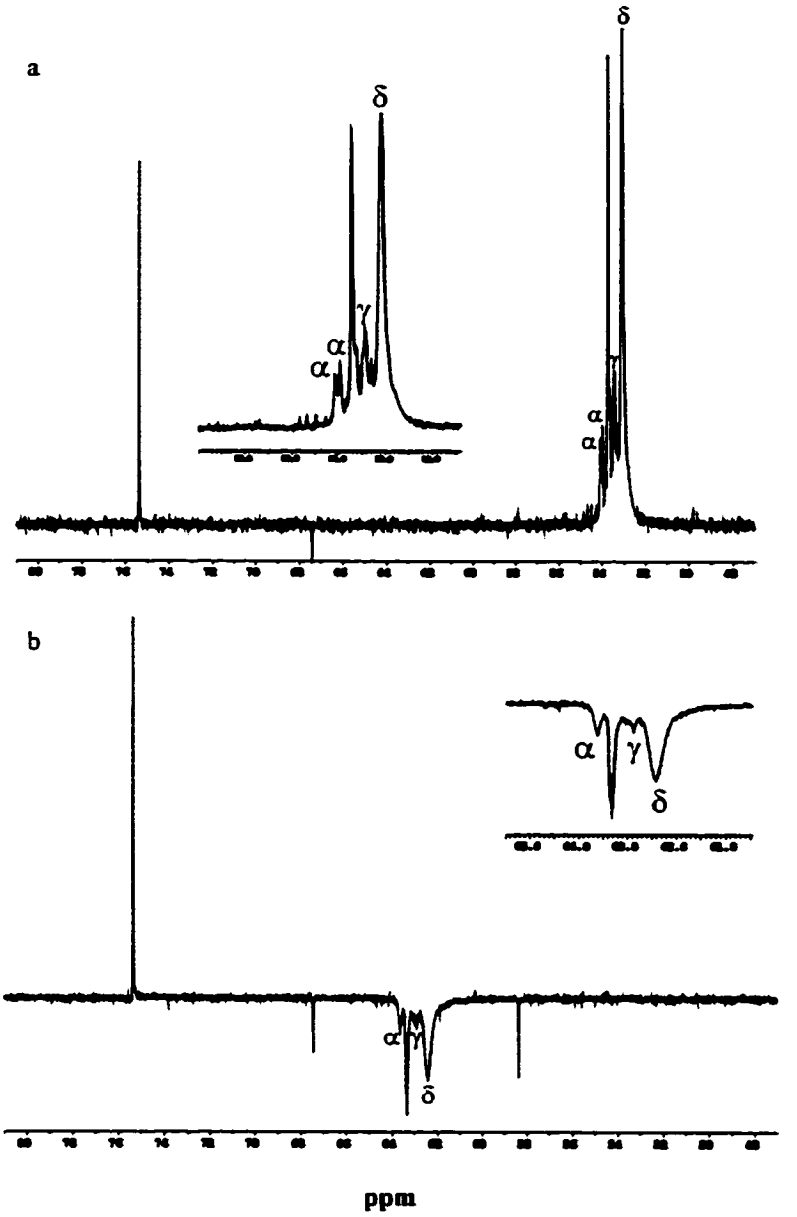


Figure 3-9. Expanded ^{13}C NMR spectra of (a) ^{13}C-methyl- and (b) ^{13}C-ethyl-esterified bovine insulin at LpH 3 (reacted *in vacuo* with ^{13}C-methanol or ^{13}C-ethanol and hydrogen chloride catalyst).

The ^{13}C-resonances corresponding to the α-, γ-, and δ-carboxyl methyl esters are: 54.08 (α), 53.97 (α), 53.73 (methylated small molecule), 53.43 (γ), 53.29 (γ), and 53.09 ppm (δ), and those corresponding to the α-, γ, and δ-carboxyl ethyl esters are: 63.61 (α), 63.37 (α), 63.32 (α), 62.87 (γ), and 62.41 ppm (δ). The resonances observed for γ-carboxyl esters may be due to methanolysis of amide groups (i.e., of Asn residues), since there are no Asp residues in the primary structure of insulin (Means and Feeney, 1971; Strickley and Anderson, 1996) or esterification of deamidated Asn or Gln residues.

to the carboxyl methyl esters are: 54.08 (α), 53.97 (α), 53.73 (small molecule methyl ester derivative), 53.43 (γ), 53.29 (γ), and 53.09 ppm (δ). There are no Asp residues in the primary structure of insulin; however, two resonances for γ-carboxyl ^{13}C-methyl esters were observed. This may result from a small amount of deamidation of the protein sample. Methanolysis of amide groups (i.e., of Asn or Gln residues) may occur in the presence of acidic methanol under the *in vacuo* esterification reaction conditions used (Chibnall, Mangan, and Rees, 1958; Means and Feeney, 1971). If methanolysis of the amide groups of Asn residues is occurring under the reaction conditions of the present study, then methanolysis of the amide groups of Gln residues may occur, but direct evidence for this is not possible since the peaks would reside in the broad resonance for the δ-carboxyl ^{13}C-methyl esters. Moreover, Strickley and Anderson (1996) have examined the mechanism and kinetics of human insulin degradation in lyophilized protein powders as a function of water content at low LpH values since studies in aqueous solution at low pH have established that decomposition proceeds through a cyclic anhydride intermediate leading to the formation of both deamidated and covalent dimer products. The rate of degradation increased with decreasing LpH (between LpH 3-5 at 35°C), and at any given water content, the LpH-rate profiles paralleled those in solution. The carboxyl group of the insulin A-chain at position 21 (A21) is a C-terminal Asn residue that is involved in rate-limiting intramolecular nucleophilic attack onto the side-chain amide carbonyl to form a reactive cyclic anhydride intermediate, which further reacts with either water (or alcohol under our reaction conditions) or an N-terminal primary amino group of another insulin molecule to generate either deamidated insulin (AspA21) or an amide-linked covalent dimer (e.g., AspA21-PheB1 or

AspA21-GlyA1). Degradation increased with water content at levels of hydration of 20 to 50%. The decomposition was accelerated by the plasticization effect (increased conformational flexibility) of sorbed water. Thus, the origin of the observed γ-carboxyl ^{13}C-methyl ester resonances in the ^{13}C-NMR spectrum of ^{13}C-methyl-esterified insulin likely arise from the esterification of the decomposition product, AspA21, at low LpH.

Insulin has two C-termini, and based on the methylation results, the resonances at 53.97 or 54.08 ppm and 53.73 ppm correspond to the α-carboxyl ^{13}C-methyl esters of C-terminal Asn and Ala residues, respectively. The extra resonance at 53.73 ppm that may correspond to an α-carboxyl ^{13}C-methyl ester may originate from the hydrolysis of the peptide backbone and the subsequent esterification of the newly formed carboxylic acid groups. Esterification of protein has been reported to occur with minor covalent modifications (Means and Feeney, 1971). The first, methanolysis of the side-chain amide groups of Asn or Gln residues has been mentioned. The second is the N→O acyl migration of Ser or Thr residues of protein. In the N→O acyl migration reaction, the N^{α}-amide bond of Ser or Thr residues may be broken by nucleophilic attack of the Ser or Thr hydroxyl group, resulting in the transformation of the amide bond into an ester bond. Under the *in vacuo* esterification reaction conditions, it is possible that the ester (unlabeled) linkages may be hydrolyzed with water (vapour) as nucleophile in acidic conditions, followed by re-esterification by acidic ^{13}C-methanol; transesterification with ^{13}C-methanol as nucleophile in the acidic conditions may also occur (Vollhardt, 1987b). The intensity of the putative α-carboxyl ^{13}C-methyl ester at 53.73 ppm is much greater than that of the other C-terminal peaks. The higher reactivity of one C-terminus may result from the manner in which the protein was prepared. Insulin was lyophilized from a

large volume, which would separate the molecules from each other to a larger extent compared to preparation from a small lyophilization volume. If one of the C-terminal residues is more exposed on the surface of the protein molecule than is the other one, then it is quite possible to increase the extent of reaction of a particular group. However, a more likely explanation of the origin of the peak at 53.73 ppm is that it arises from esterification or methylation of a small molecule (which may be present as a counterion).

Figure 3-9b shows the ^{13}C-NMR spectrum of *in vacuo* ^{13}C-ethyl-esterified insulin at LpH 3. The spectrum illustrates the three types of ^{13}C-ethyl esters of protein α-, γ-, and δ-carboxyl groups. The chemical shifts of the ^{13}C-resonances corresponding to the carboxyl ethyl esters are: 63.61 (α), 63.37 (α), 63.32 (α), 62.87 (γ), and 62.41 ppm (δ). As was the case with one of the α-carboxyl ^{13}C-methyl ester derivatives of insulin, the intensity of the resonances of the two α-carboxyl ^{13}C-ethyl esters at 63.37 and 63.32 ppm is greater than the resonance of the α-carboxyl ^{13}C-ethyl ester at 63.61 ppm. A similar explanation as the one given for the α-carboxyl ^{13}C-methyl ester derivative of insulin can be put forth for the α-carboxyl ^{13}C-ethyl ester derivatives of insulin. Moreover, as was the case for ^{13}C-methyl-esterified insulin, the resonance observed for the γ-carboxyl ^{13}C-ethyl ester at 62.87 ppm is probably the result of methanolysis of amide groups (i.e., of Asn residues), since there are no Asp residues in the primary structure of insulin (Means and Feeney, 1971).

3.3.3 α-Chymotrypsin

Figure 3-10a shows the ^{13}C-NMR spectrum of *in vacuo* ^{13}C-methyl-esterified α-chymotrypsin at LpH 3, and clearly illustrates the three types of ^{13}C-methyl esters of protein α-, γ-, and δ-carboxyl groups. The chemical shifts of the ^{13}C-resonances

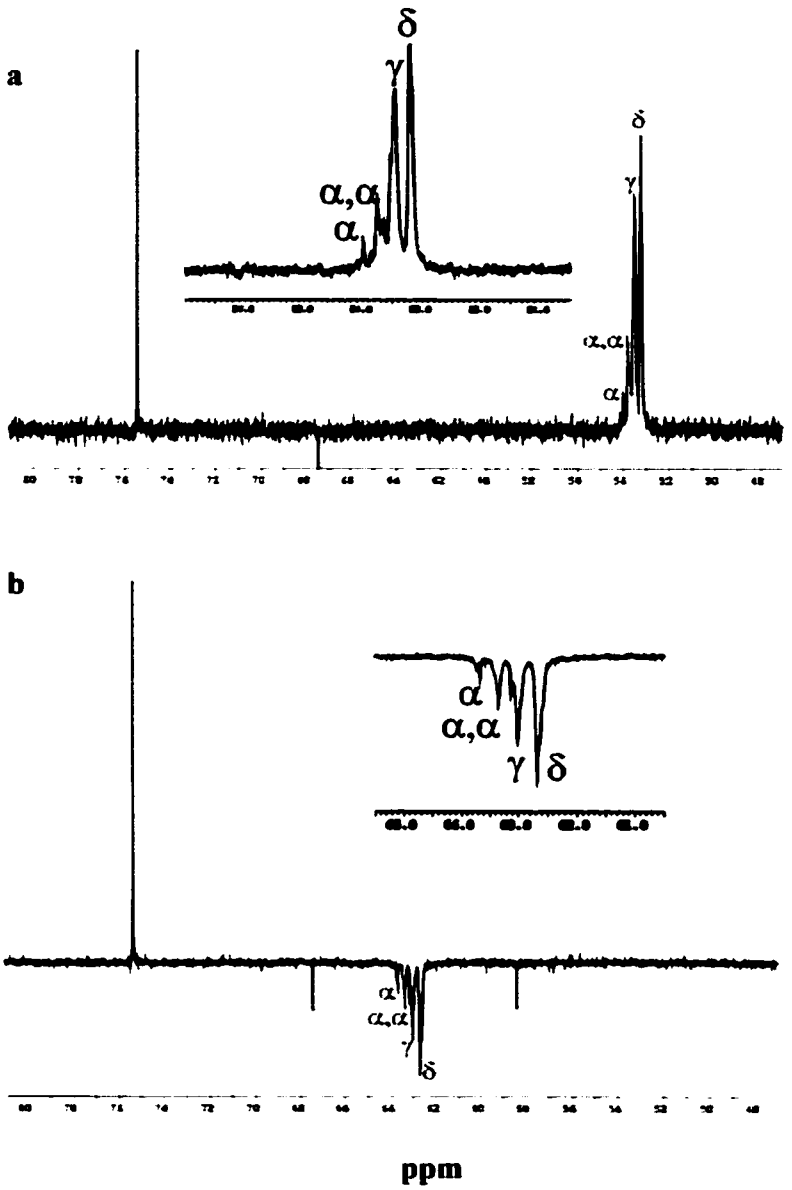


Figure 3-10. Expanded ^{13}C NMR spectra of (a) ^{13}C-methyl- and (b) ^{13}C-ethyl-esterified α-chymotrypsin at LpH 3 (reacted *in vacuo* with ^{13}C-methanol or ^{13}C-ethanol and hydrogen chloride catalyst).

The ^{13}C-resonances corresponding to the α-, γ-, and δ-carboxyl methyl esters are: 53.98 (α), 53.76 (α), 53.69 (α), 53.33 (γ), 53.47 (γ), and 53.21 ppm (δ), and those corresponding to the α-, γ-, and δ-carboxyl ethyl esters are: 63.73 (α), 63.67 (α), 63.35 (α), 63.12 (γ), 63.01 (γ), 62.67 (δ), and 62.63 ppm (δ).

corresponding to the carboxyl methyl esters are: 53.98 (α), 53.76 (α), 53.69 (α), 53.53 (γ), 53.47 (γ), and 53.21 ppm (δ). The α-carboxyl ^{13}C-methyl ester resonance at 53.98 ppm corresponds to the C-terminal Asn-245 residue (see esterified insulin description above) and is downfield to the other two (or more) α-carboxyl ^{13}C-methyl ester resonances. More than two of the latter resonances are observed between 53.60-53.85 ppm (the peak picking function in the NMR program did not pick all the observed peaks), and are likely the result of autoproteolysis. Comparison with α-chymotrypsin ^{13}C-methylated with ^{13}C-iodomethane, two of the C-terminal resonances, Leu-13 and Tyr-146, are observed in the range of 53.72-53.77 ppm, which corresponds to those observed with α-chymotrypsin ^{13}C-methyl-esterified with ^{13}C-methanol/HCl (53.76 and 53.69 ppm).

Figure 3-10b shows the ^{13}C-NMR spectrum of *in vacuo* ^{13}C-ethyl-esterified α-chymotrypsin at LpH 3. The spectrum illustrates the three types of ^{13}C-ethyl esters of protein α-, γ-, and δ-carboxyl groups. The chemical shifts of the ^{13}C-resonances corresponding to the carboxyl ethyl esters are: 63.73 (α), 63.67 (α), 63.35 (α), 63.12 (γ), 63.01 (γ), 62.67 (δ), and 62.63 ppm (δ). The resonance at 63.73 or 63.67 ppm corresponds to the α-carboxyl ^{13}C-ethyl ester of Asn-245. It is possible that both resonances correspond to that for Asn-245 if this C-terminal residue is in a different environment in one molecule compared to another. The resonance at 63.35 ppm is broad and larger than that for Asn-245 at 63.67-63.73 ppm, and corresponds to the resonances for the other two C-terminal α-carboxyl ^{13}C-ethyl esters of α-chymotrypsin (Leu-13 and Tyr-146). Chymotrypsin has 9 Asp and 5 Glu residues, and although there is a small amount of broadening of the γ- and δ-carboxyl ester ^{13}C-resonances, it appears that all the

carboxyl resonances in each group have very similar chemical shifts regardless of the alcohol used in the reaction.

Table 3-4 summarizes the chemical shifts of α-, γ-, and δ-carboxyl ^{13}C-methyl or ^{13}C-ethyl ester resonances for the *in vacuo* ^{13}C-esterified proteins studied. The data clearly illustrates that the resonances for α-, γ- and δ-carboxyl ^{13}C-methyl or ^{13}C-ethyl esters reside in specific and distinct regions of the ^{13}C-NMR spectrum. The chemical shifts of the resonances among the esterified α-carboxyl groups showed a much greater variation than among those of the esterified γ- or δ-carboxyl groups, indicating that the chemical shift of the α-carboxyl methyl ester resonance in proteins is much more sensitive to its environment than that for either of the γ- or δ-carboxyl ^{13}C-methyl esters. The number of α-carboxyl resonances observed for the methyl and ethyl esters corresponded to the number of α-carboxyl groups known to be present in the protein. For ribonuclease (single chain), one resonance for the α-carboxyl ^{13}C-methyl ester is present (Figure 3-8); and for α-chymotrypsin (three chains), multiple resonances for α-carboxyl ^{13}C-methyl esters are observed (Figure 3-10). For insulin (two chains), it appears that two resonances for the α-carboxyl ^{13}C-methyl esters were observed; however, a sharp peak at 53.73 ppm with an intensity six times that of the latter resonances was evident in the α-carboxyl ^{13}C-methyl ester region, which is likely that for a small molecule present as a counterion in the protein matrix, perhaps phosphate; the intensity of the peak in question is too large to arise from an α-carboxyl ^{13}C-methyl ester. Esterified γ- and δ-carboxyl groups in proteins had a significant difference in their chemical shifts (approximately 0.3 ppm) and could be easily distinguished (Table 3-4). However, the

Table 3-4. Chemical Shifts of α-, γ-, δ-Carboxyl ^{13}C-Methyl or ^{13}C-Ethyl Ester Resonances in *In Vacuo* ^{13}C-Esterified Proteins

^{13}C-methyl-esterified protein	Chemical shifts of $C_\alpha OO^{13}Me$ (ppm)	Chemical shifts of $C_\gamma OO^{13}Me$ (ppm)	Chemical shifts of $C_\delta OO^{13}Me$ (ppm)
ribonuclease A	53.74	53.46, 53.41	53.18
bovine insulin	54.08 or 53.97, 53.73	53.43, 53.29	53.09
α-chymotrypsin	53.98, 53.76, 53.69	53.53, 53.47	53.21
^{13}C-ethyl-esterified protein	**Chemical shifts of $C_\alpha OO^{13}Et$ (ppm)**	**Chemical shifts of $C_\gamma OO^{13}Et$ (ppm)**	**Chemical shifts of $C_\delta OO^{13}Et$ (ppm)**
ribonuclease A	63.37	63.02, 62.79	62.64, 62.45
bovine insulin	63.61, 63.37, 63.32	62.87	62.41
α-chymotrypsin	63.73, 63.67, 63.35	63.12, 63.01	62.67, 62.63

individual esterified carboxyl groups within each class showed a much smaller variation in their chemical shifts.

Comparison of the ^{13}C-NMR spectra of any of the methyl- or ethyl-esterified proteins studied suggests that the extent of ethyl-esterification is less than that for methyl-esterification. However, this conclusion cannot be made since the enriched carbon atom is a methylene group (in an ethanol molecule) that has only two protons from which polarization transfer (DEPT-135) can occur (to enhance the sensitivity of the carbon-13 resonance) as opposed to the three protons of a ^{13}C-methyl group in a methanol molecule. Moreover, relaxation of carbon-13 atoms in methylene and methyl groups is different. Hence, the extent of esterification with either alcohol must be determined by other means. The approach employed in this work was the difference titration methodology and it was applied to methyl- and ethyl-esterified ribonuclease A (Table 3-5).

3.4 Determination of the Extent of Methylation or Esterification of Ribonuclease A Carboxyl Groups

3.4.1 Extent of Methylation of Ribonuclease A Carboxylate Salts with ^{13}C-Iodomethane

Figure 3-11 shows the ^{13}C-NMR spectra of ^{13}C-methylated ribonuclease A for an LpH course experiment. The extent of derivatization is directly related to the pH memory effect (Zaks and Klibanov, 1988). The pH of the aqueous solution from which the protein was lyophilized determines the ionization state of a functional group. It is the ionization state of a group, its intrinsic nucleophilicity, and its microenvironment in the protein molecule (Kaplan *et al.*, 1971; Young and Kaplan, 1989), which determines or affects the relative reactivity of the various functional groups of proteins in solution.

These factors appear to be retained in the lyophilized state, as evidenced by the results obtained for the LpH course experiment of ribonuclease A reacted *in vacuo* with ^{13}C-iodomethane. It was observed that the degree of carbon-13 incorporation into protein increased with LpH. From the ^{13}C-NMR spectra, it was possible to semi-quantitatively compare relative peak intensities for the resonances corresponding to the carboxyl ^{13}C-methyl ester and ^{13}C-trimethyl amino derivatives of N-terminal α-amino (Lys for ribonuclease A) and lysyl ε-amino groups. The pK_a values of carboxyl groups are low (3.5-4.5) compared to those for amino groups (6.8-11.1) (Creighton, 1993b). It was observed from the peak intensities of their derivatives that carboxyl groups appeared to react with ^{13}C-iodomethane to their maximum extent at LpH 6-7, while lysyl ε-amino groups apparently reacted to their full extent at LpH 10, as expected from the respective pK_a values of these groups. Moreover, the intrinsic nucleophilicity of carboxyl groups is less than that for amino groups, and the latter groups have been shown to react almost fully by Taralp (Doctoral Thesis, 1997b). It was shown that nonaqueous *in vacuo* ^{13}C-methylated insulin (24 h, T=75°C, LpH=10) tested weakly ninhydrin positive and the aqueous (control) methylated insulin samples (24 h, T=75°C, pH=10) tested ninhydrin negative.

3.4.1.1 Estimation of the Extent of *In Vacuo* Methylation of Ribonuclease A Carboxylate Salts by Difference Titration

There is no convenient and rapid method available for the estimation of the ester content in protein. The difference titrimetric method used is time consuming and gives a variation of 10-20% with four or more determinations. The data show (Table 3-5, Section 3.4.1) an increase in the degree of alkylation of carboxyl groups with LpH; however, the extent of reaction is only approximately 40% complete at LpH 7.

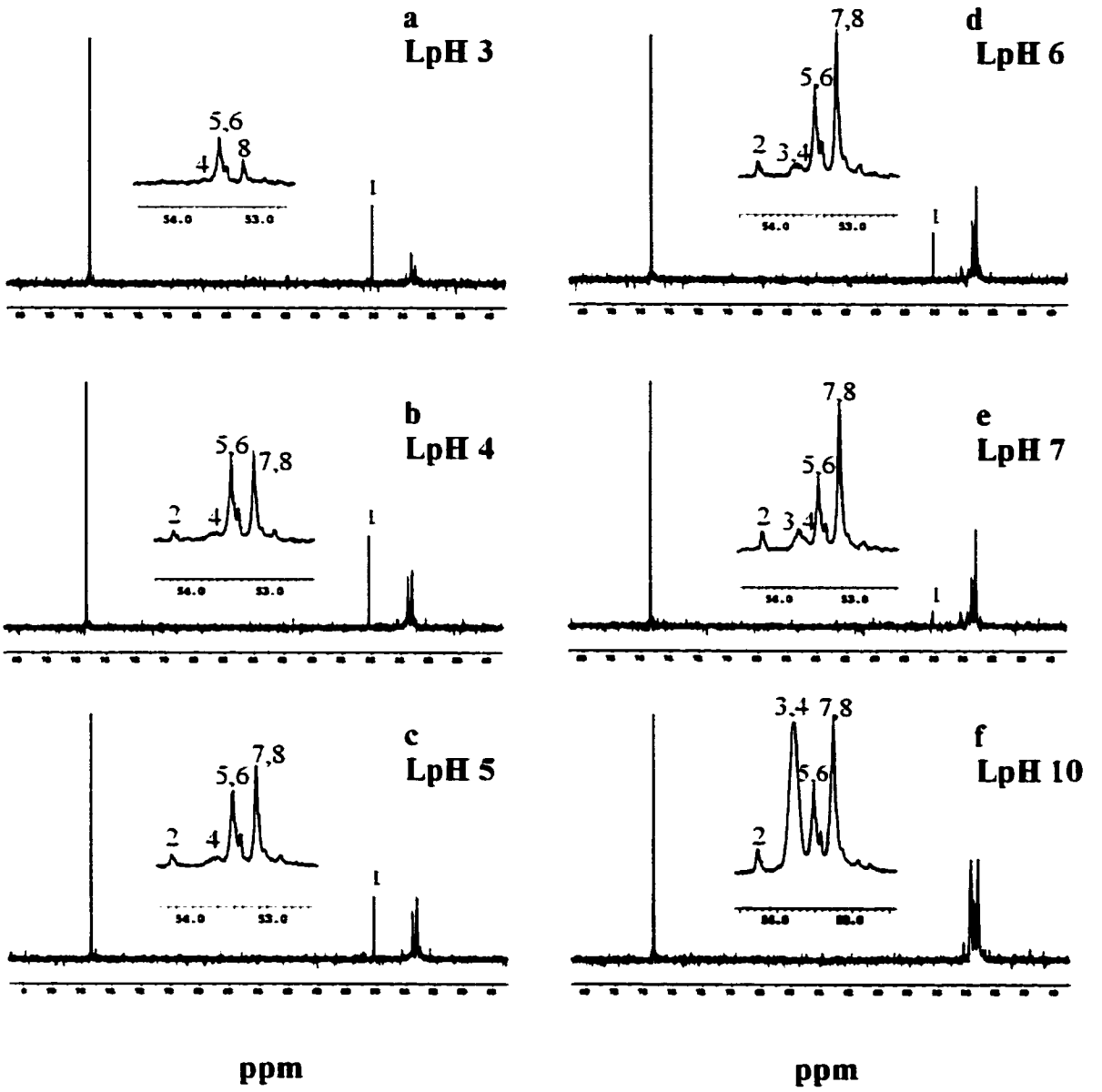


Figure 3-11. Expanded ^{13}C NMR spectra of ^{13}C-methylated ribonuclease A (10 mg, no buffer) at various LpH values (LpH 3-10; reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h).

Peak resonances correspond to the ^{13}C-enriched methyl groups in:
(1) unidentified small molecule (56.21 ppm); (2) unidentified α-$^{+}NMe_3$ impurity (54.23 ppm); (3) $Me_3N^{\varepsilon+}$-Lys (53.76 ppm); (4) $C_\alpha OOMe$ (53.71 ppm); (5) $C_\gamma OOMe$ (53.56 ppm); (6) $C_\gamma OOMe$ (53.48 ppm); (7) $Me_3N^{\alpha+}$-Lys (53.24 ppm); (8) $C_\delta OOMe$ (53.23 ppm).

Unfortunately, the difference in the extent of methylation of carboxyl groups is not easily observable from the ^{13}C-NMR spectra for ^{13}C-methylated ribonuclease A at LpH 7 and 10, due to interfering resonances from N^{ϵ}-trimethyl amino groups of lysyl residues. Moreover, the difference titration data for the ^{13}C-methylated ribonuclease A at LpH 10 gave a value of 58%, which is much higher than that for LpH 7. It is reasonable to suggest that this value is abnormally high on the basis of the pK_a values of carboxyl groups in protein (pK_a 3.5-4.5), as these groups should have reacted fully at LpH 7. However, it is possible that the pK_a value of a carboxyl group is perturbed by its microenvironment (e.g., pK_a may increase if carboxylic acid group is in a hydrophobic environment, in which case the group would be reluctant to give up its proton and would probably be in the hydrophobic core of the protein molecule and not exposed to reagent). The high value for the extent of reaction of ^{13}C-methylated ribonuclease A at LpH 10 may be due to the protein molecules unfolding at the high pH prior to lyophilization. Buried carboxylate salts could thus be exposed, which can react potentially with reagent to lead to an increase in the extent of reaction. Moreover, at LpH 10 there are deprotonated amino groups present that can buffer the hydroiodic acid that is liberated from the methylation reaction, thus promoting the S_N2 reaction of carboxylate salt groups with iodomethane (March, 1985c) for a longer duration of time, leading to a higher extent of reaction. Furthermore, it seems improbable that the extent of reaction is as high as 58% for the methylation of ribonuclease A at LpH 10, since this value is close to that obtained for the esterification reaction of this protein with ^{13}C-methanol/HCl at LpH 3 (see Table 3-5, Section 3.4.1). The ^{13}C-NMR spectrum of the latter sample is shown in Figure 3-8a, and clearly shows the extent of reaction of carboxyl groups to be 88%

greater (based on integrated peak areas) than that for the methylation reaction (cf. Figure 3-11). It should be noted that some variability in the extent of esterification with ^{13}C-methanol/HCl at LpH 3 was observed (see below); however, the latter reaction always revealed, semi-quantitatively by ^{13}C-NMR spectroscopy, a larger extent of reaction (by 88%).

3.4.1.2 Effect of the Addition of Small Amounts of Water on the Extent of *In Vacuo* ^{13}C-Methylation Reaction of Ribonuclease A

One reason for the variability in the extent of reaction could be the amount of water present in the lyophilized powder. Esterification is catalyzed by acids (not bases) via bimolecular acyl cleavage (AAC2), and sometimes by unimolecular acyl or alkyl cleavage (AAC1 and AAL1, respectively) mechanisms (March, 1985a). Since esterification and hydrolysis of esters are reversible reactions they have the same mechanisms. An experiment was conducted with increasing amounts of distilled water (0, 1, 2, 5 μL) added to the reaction vessels for the ^{13}C-methanol/HCl (25 μL:10 μL) reaction *in vacuo* with ribonuclease A at LpH 3. The results obtained showed a progressive decrease in the extent of esterification of ribonuclease A (Figure 3-12), as expected since water would shift the equilibrium of the reaction toward hydrolysis and sway it from the synthetic route (esterification). A similar experiment was conducted for the alkylation reaction using ^{13}C-iodomethane and ribonuclease A at LpH 6 (Figure 3-13). As expected, a progressive decrease in the extent of methylation with added distilled water to the reaction vessel was observed. Interestingly, with increasing amounts of added distilled water the extent of reaction with α- or ε-amino groups decreased. This lends evidence to the hypothesis that water binds to charged and polar functional groups first in the lyophilized protein (Rupley, Gratton, and Careri, 1983).

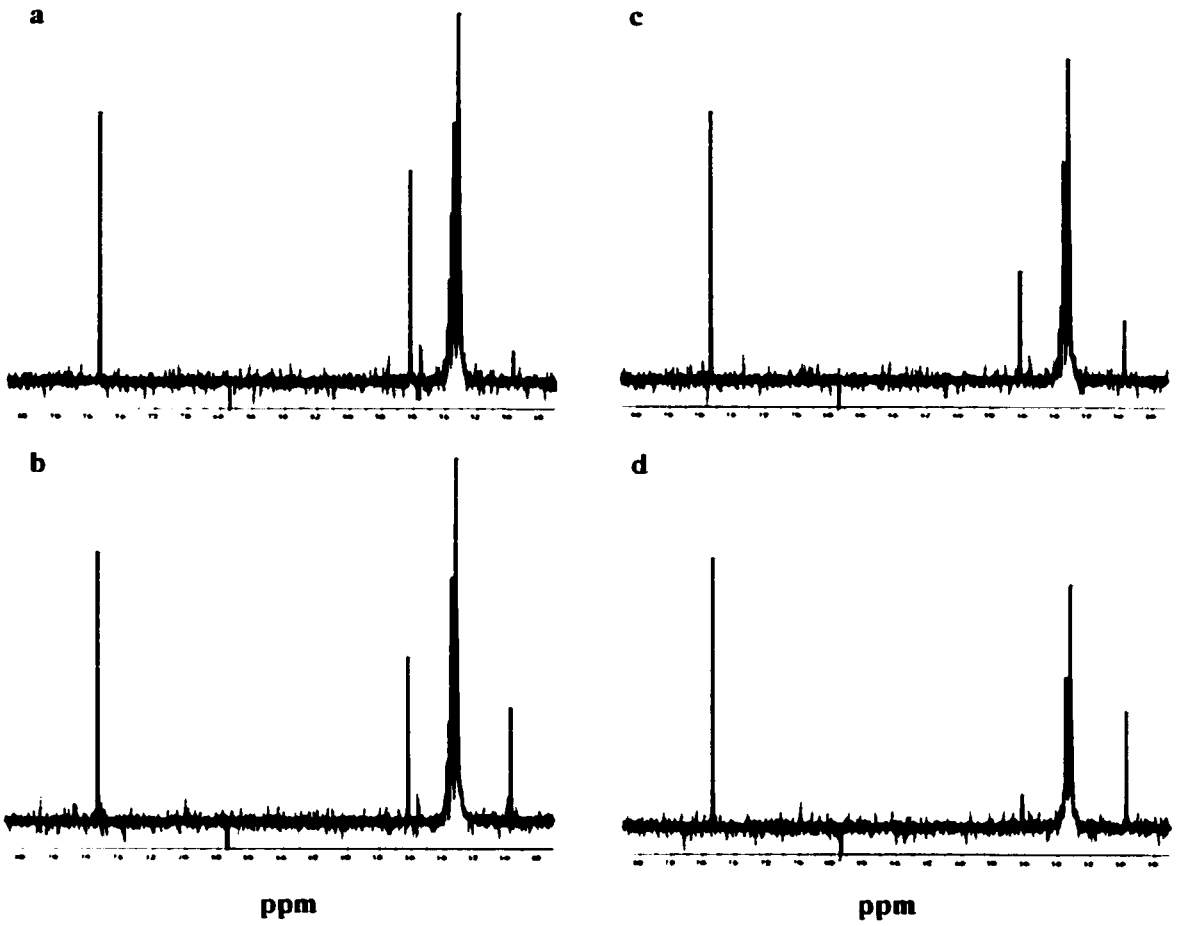


Figure 3-12. Effect of addition of water on the extent of *in vacuo* ^{13}C-methyl-esterification reaction of ribonuclease A (20 mg) at LpH 6.

Increasing amounts of distilled water: (a) 0 μL, (b) 1 μL, (c) 2 μL, (d) 5 μL were added to the reaction vessels after addition of the ^{13}C-methanol/HCl (25 μL:20 μL) esterification reagents.

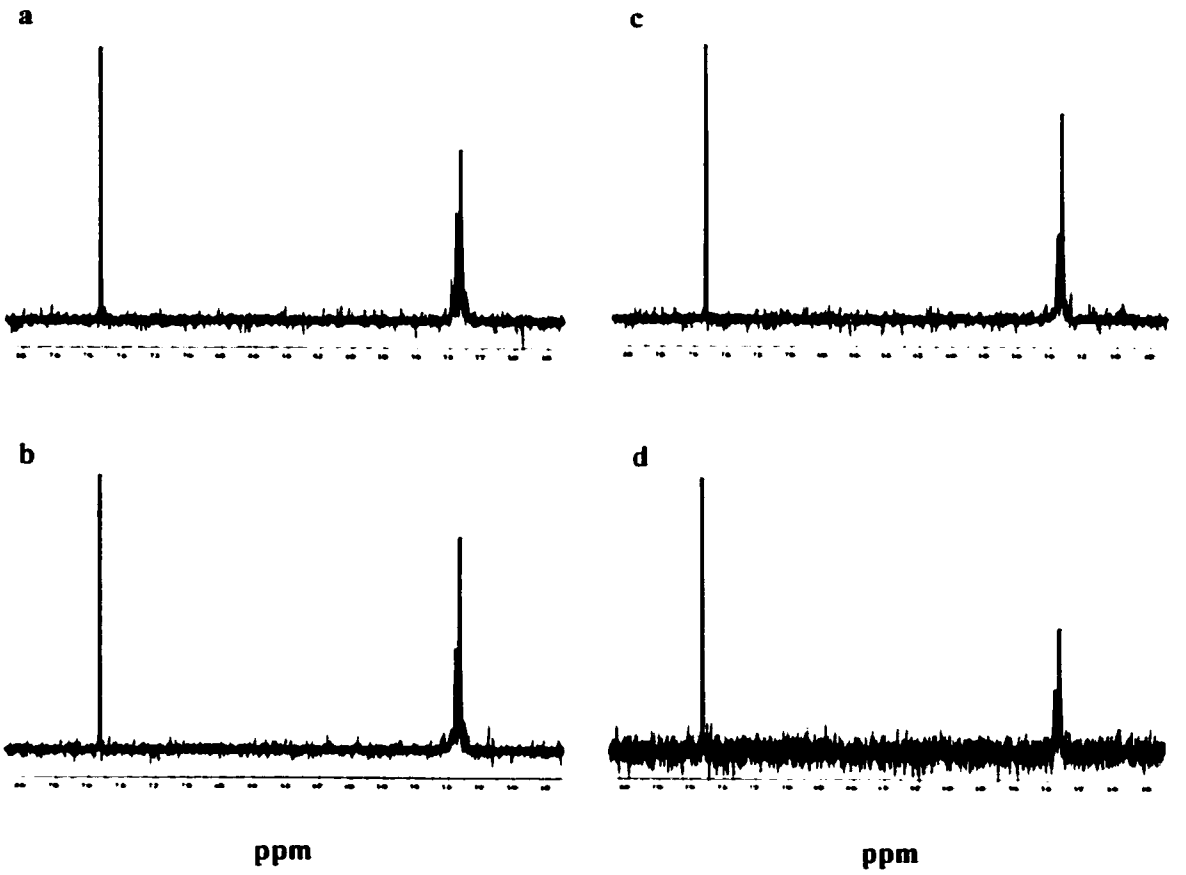


Figure 3-13. Effect of addition of water on the extent of *in vacuo* ^{13}C-methylation reaction of ribonuclease A (20 mg) at LpH 6.

Increasing amounts of distilled water: (a) 0 μL, (b) 1 μL, (c) 2 μL, (d) 5 μL were added to the reaction vessels after addition of ^{13}C-iodomethane (25 μL).

Additional water in the system may also react with iodomethane to form methanol and hydroiodic acid. The addition of 5 μL distilled water resulted in what appears to be the preferential derivatization of carboxyl groups with ^{13}C-iodomethane, which would support the latter possibility.

3.4.1.3 Effect of Reaction Conditions (LpH, time, temperature, presence or absence of buffer, reagent amount) on the Extent of *In Vacuo* ^{13}C-Methylation Reaction of Ribonuclease A

It is noteworthy that the results for the extent of methylation were for unbuffered lyophilized ribonuclease A. It has been observed previously by Taralp (Doctoral Thesis, 1997), and confirmed by this author, that for ^{13}C-methylated ribonuclease A at LpH 10, the resonance of N^{ε}-trimethyl amino groups of lysyl residues is very large when buffer (100 mM sodium metaborate) is present cf. when buffer is absent from the lyophilized powder (cf. Figure 3-14 (d) and (j)). The reason for this increase lies in the mechanism of the methylation reaction. The trimethylammonium derivative of lysyl ε-amino groups is presumably formed via a series of deprotonation and alkylation steps (Vollhardt, 1987a). The liberated hydroiodic acid must be transferred to and neutralized by lyophilized buffer or base for the methylation to proceed as observed. The high temperature of the reaction or the volatility of hydroiodic acid under evacuated conditions may help promote such a transfer. However, if buffer is absent from the lyophilized powder, the extent of methylation is significantly less, i.e., the N^{ε}-amino groups of lysyl residues react less (Figure 3-14 (d)). This is due probably to the decrease in LpH of the protein by volatile hydroiodic acid; in this case, the deprotonated amino groups would accept the proton from hydroiodic acid, rendering them unreactive with iodomethane. In order to determine whether buffering, higher temperature, or a longer reaction time could

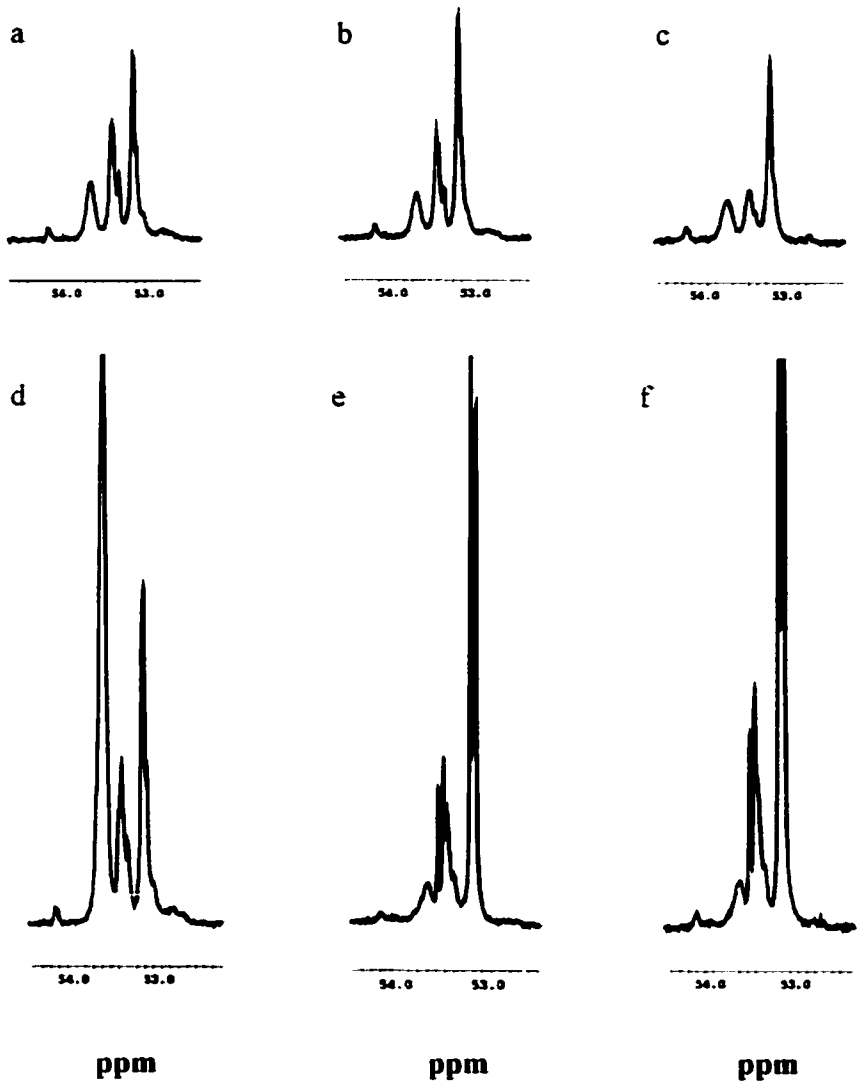


Figure 3-14(a)-(f). Effect of conditions (LpH, time, temperature, presence or absence of buffer, reagent amount) on the extent of *in vacuo* ^{13}C-methylation of ribonuclease A (20 mg).

Typical reaction conditions: 25 μL ^{13}C-iodomethane, 75°C, 24 h.
Samples without buffer (NB): (a) LpH 6, 24 h, 75°C; (b) LpH 6, 72 h, 75°C; (c) LpH 6, 72 h, 100°C; (d) LpH 10, 24 h, 75°C. **Samples with sodium phosphate buffer (B):** (e) LpH 6, B, 24 h, 75°C; (f) LpH 6, B, 72 h, 75°C.

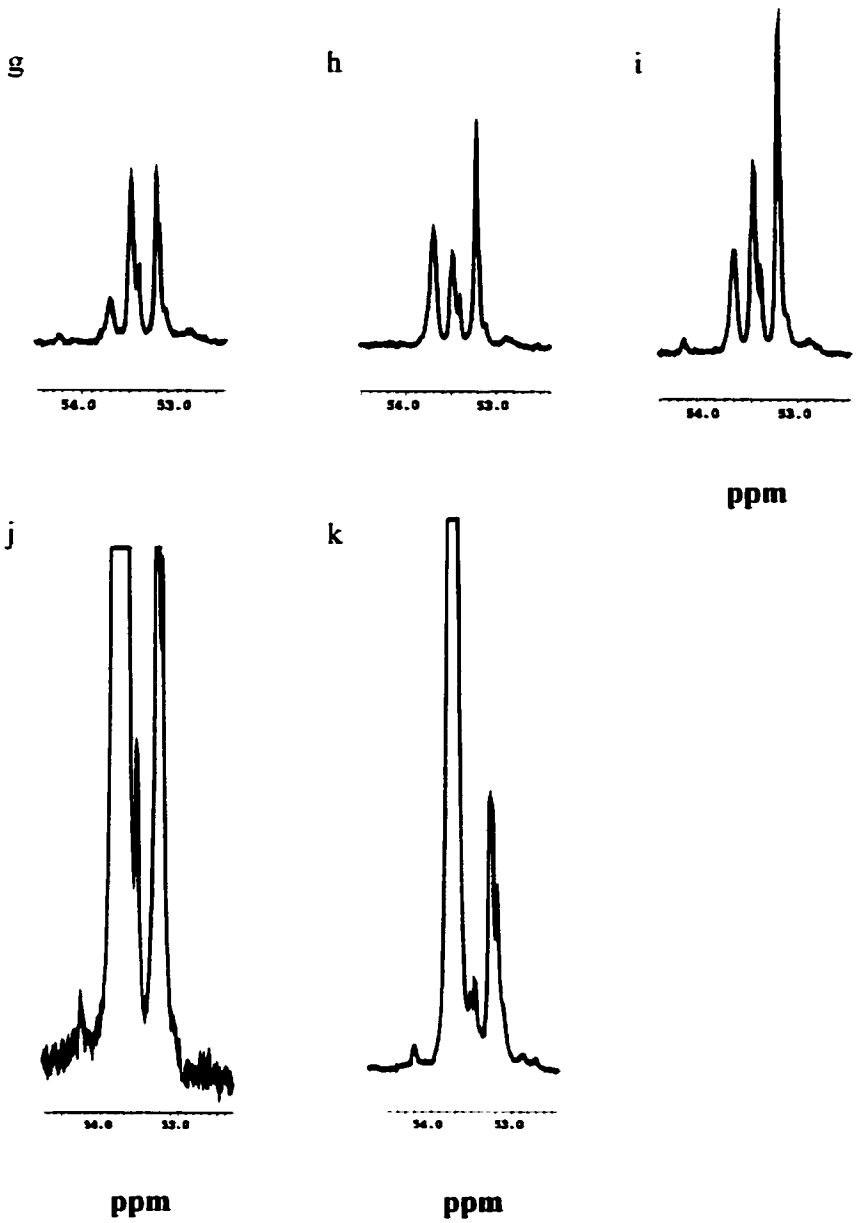


Figure 3-14(g)-(j). Effect of conditions on the extent of *in vacuo* ^{13}C-methylation of ribonuclease A (20 mg).

Typical reaction conditions: 25 μL ^{13}C-iodomethane, 75°C, 24 h. **Dialyzed samples** (against 0.001 M HCl): (g) sample (f) dialyzed; (h) LpH 6, NB, 24 h, 75°C. **Reagent amount increased:** (i) 50 μL ^{13}C-iodomethane, LpH 6, NB, 24 h, 75°C. **Samples with sodium metaborate buffer (B):** (j) LpH 10, B, 24 h, 75°C. **Dialyzed sample:** (k) sample (j) dialyzed.

increase the extent of methylation of protein carboxylate salt groups, *in vacuo* methylations with ^{13}C-iodomethane were carried out on ribonuclease A at LpH 6 and LpH 10 with and without buffer, at 75°C and 100°C for 24 and 72 h.

It was observed for ribonuclease A, LpH 6 without buffer, ^{13}C-methylated with ^{13}C-iodomethane (72 h, 75°C) that there was very little difference in reactivity of the functional groups of interest cf. the sample reacted for 24 h at 75°C (cf. Figure 3-14 (a) and (b)). A slight increase in the N^{α}-trimethyl Lys α-amino resonance at 53.2 ppm, and a slight decrease in the N^{ε}-trimethyl Lys ε-amino resonance at 53.7 ppm occurred, while the resonances for the γ-carboxyl ^{13}C-methyl esters at 53.4-53.5 ppm remained similar to the 24 h methylation at 75°C. Reaction at 100°C for 72 h showed a slight decrease in all relevant resonances (Figure 3-14 (c)). Comparison of ^{13}C-methylated ribonuclease A, LpH 10 (without buffer, 24 h, 75°C; Figure 3-14 (d)) with that at LpH 6 revealed very little difference in the extent of methylation for γ-carboxyl ^{13}C-methyl ester resonances at 53.4-53.5 ppm. There was the expected increase in peak intensities for N^{α}- and N^{ε}-trimethyl amino groups of lysyl residues at 53.2 ppm and 53.7 ppm, respectively; however, since the δ-carboxyl ^{13}C-methyl ester resonances at 53.2 ppm overlap that of the N^{α}-trimethyl amino group of the lysyl N-terminal residue, it is difficult to comment on the extent of reaction for those functional groups. ^{13}C-Methylation of ribonuclease A, LpH 10 for 72 h did not improve the extent of reaction of carboxyl or amino groups, and actually diminished it slightly (spectrum not shown).

Ribonuclease A, LpH 6 ^{13}C-methylated in the presence of phosphate buffer for 24 h at 75°C showed a lesser extent of reaction for all peaks of interest (Figure 3-14 (e)); whereas, a longer reaction time under the same reaction conditions revealed an increase

in the extent of reaction of γ-carboxyl groups with ^{13}C-iodomethane (Figure 3-14 (f)). Phosphate buffer reacted with ^{13}C-iodomethane to yield methyl ester derivatives, which were evidenced as a doublet at 53.57, 53.51 ppm (trimethylphosphate) prior to hydroxylamine and/or base treatments and at 53.21, 53.15 ppm (dimethylphosphate) after these treatments. The chemical shifts of the doublet for these proposed derivatives were assigned based on those for dimethyl phosphite (52.10, 52.02 ppm) and trimethyl phosphate (54.37, 54.29 ppm); the observed J couplings between ^{13}C and ^{31}P (I = ½) appear to be comparable to those of the latter model compounds. These resonances obscured the analysis of the carboxyl methyl esters, and were removed from the ^{13}C-NMR spectra by dialyzing the respective samples. What seemed contradictory to the above, the ^{13}C-NMR spectrum of dialyzed ^{13}C-methylated ribonuclease A, LpH 6 with sodium phosphate buffer (75°C, 72 h) was similar to that for the sample reacted under the same reaction conditions but for 24 h. The former dialyzed sample is shown in Figure 3-14 (g). Comparison of this sample with that for ^{13}C-methylated ribonuclease A, LpH 6 without buffer (24 h, 75°C) (Figure 3-14 (h)), dialyzed before the methylation reaction, and after the acquisition of the first ^{13}C-NMR spectrum, showed an increase in the extent of reaction of γ-carboxyl groups, as shown by an increase in the γ-carboxyl ^{13}C-methyl ester resonances at 53.4-53.5 ppm (cf. Figure 3-14 (g) and (h)). Moreover, the increase in the extent of reaction of γ-carboxyl groups with ^{13}C-iodomethane for ribonuclease A, LpH 6, lyophilized in the presence of phosphate buffer, approached that observed for ^{13}C-methylated ribonuclease A, LpH 6 without buffer (24 h, 75°C; dialyzed after the first NMR analysis) (Figure 3-14 (h)), using 50 μL ^{13}C-iodomethane cf. the usual 25 μL (see below discussion) (cf. Figure 3-14 (g) and (i)) and approached that

observed for ^{13}C-methylated ribonuclease A, LpH 10 (without buffer; 25 μL ^{13}C-iodomethane; cf. Figure 3-14 (g) and (d)). However, the extent of trimethylation of N^{α}- and N^{ε}-amino groups of lysyl residues did not approach that observed for the sample methylated with 50 μL ^{13}C-iodomethane.

^{13}C-Methylation of ribonuclease A, LpH 10 in the presence of sodium metaborate buffer (24 h, 75°C; Figure 3-14 (j)) resulted in a large increase in the N^{ε}-trimethyl amino resonance of lysyl residues and an apparent increase in the N^{α}-trimethyl amino resonance of the N-terminal lysyl residue. The latter resonance was difficult to interpret, since the buffer ^{13}C-methyl ester peaks overlap it. It was not possible to make a conclusion with regard to the extent of reaction of γ-carboxyl groups, since some hydrolysis of the labile ^{13}C-methyl ester groups occurred during ^{13}C-NMR sample preparation. ^{13}C-Methylation of ribonuclease A, LpH 10 reacted under similar condition for 72 h at 75°C did not offer any additional information. Hence, it is very important that care be taken during sample preparation of higher LpH samples. When acidic D_2O or urea/D_2O is not used, hydrolysis of ^{13}C-methyl esters is evidenced by an increase in the resonance for ^{13}C-methanol at 49.8 ppm, and the resonances for the carboxyl ^{13}C-methyl esters are weaker than they initially were. The ^{13}C-methylated ribonuclease A, LpH 10 (with sodium metaborate) sample was dialyzed and is shown in Figure 3-14 (k). Unfortunately, the sample was spilt accidentally and half of it was lost; additional D_2O was added necessarily to acquire the ^{13}C-NMR spectrum; however, the internal standard was present and the relative peak intensities would be the same as for the more concentrated (regular) sample; but, the trade-off was that the signal-to-noise ratio is poorer. The results indicate that there was an increase in the resonances for the γ-carboxyl ^{13}C-methyl esters.

Moreover, on the same lines of determining if any other obvious variables were the cause of the seemingly low extents of the methylation reaction, the amount of reagent added was also examined. The usual amount of ^{13}C-iodomethane used for 20 mg ribonuclease A was 25 μL (4×10^{-4} μmol), which is well in excess to the amount of reactable functional groups (e.g., 60-70 μmol reactable groups = 6×10^{-5} μmol) in the protein molecules on a mole-to-mole basis. However, if buffer from the protein preparation was present in undialyzed samples or trapped in the protein matrix in dialyzed samples, then it is quite possible that insufficient reagent was being used. Ribonuclease A (20 mg) was reacted *in vacuo* with 50 μL ^{13}C-iodomethane; a small, but significant, increase in the extent of reaction was indicated by the ^{13}C-NMR spectrum. Although difference titrations were not carried out for samples obtained under these reaction conditions, the ^{13}C-NMR spectroscopic data revealed an increase of approximately 20% in reaction extent for all reactable functional groups compared to ribonuclease A reacted with 25 μL ^{13}C-iodomethane (cf. Figure 3-14 (a) and (i); based on integrated peak areas).

3.4.1.4 Effect of Reaction Conditions (external base, reagent amount) on the Extent of *In Vacuo* ^{13}C-Methylation Reaction of Unbuffered Ribonuclease A

In an attempt to maintain the LpH of the lyophilized protein at the desired value, experiments were carried out with base placed externally in the larger borosilicate glass tube reaction vessel and the smaller borosilicate glass culture tube in which the lyophilized protein was contained was separated from the base powder with a small piece (1 cm) of glass rod. The base was used to react with hydroiodic acid (HI), and thus to prevent the inhibition of the reaction by the latter. Crushed sodium hydroxide or sodium hydrogenphosphate heptahydrate (crystalline or lyophilized) were the bases tested.

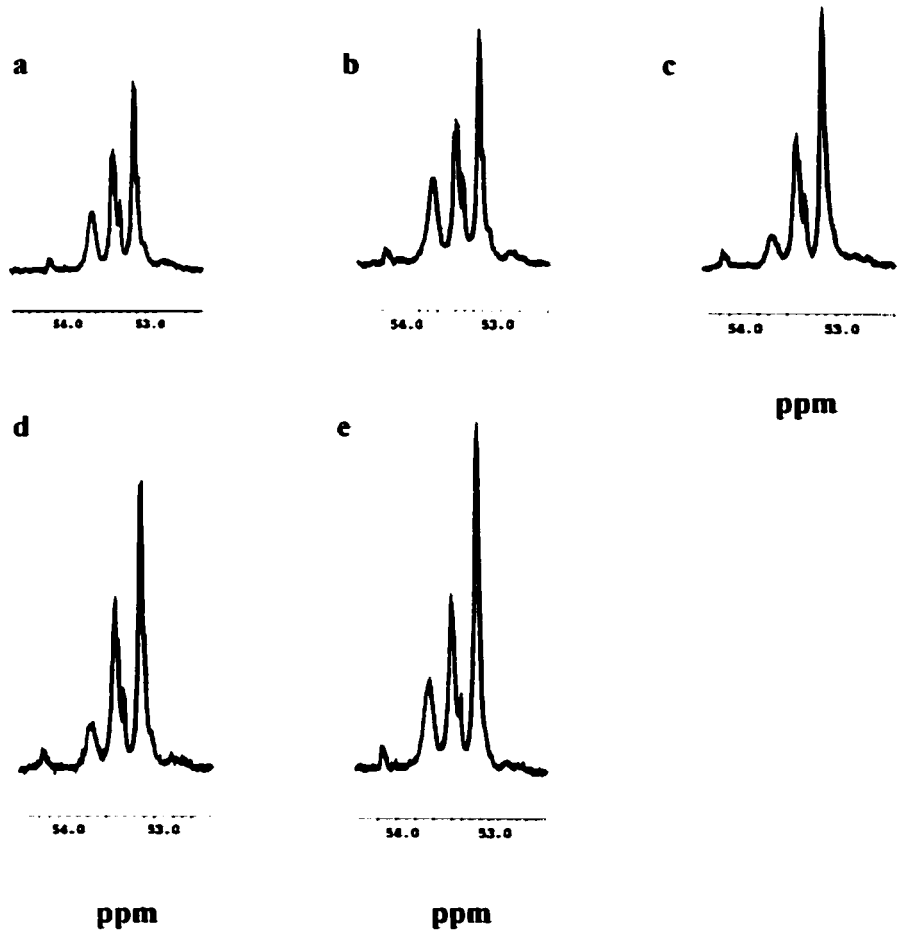


Figure 3-15. Effect of conditions (external base, reagent amount) on the extent of *in vacuo* ^{13}C-methylation of unbuffered ribonuclease A (20 mg) at LpH 6.

Typical reaction conditions: 25 μL ^{13}C-iodomethane, 75°C, 24 h. (a) control; (b) NaOH external; (c) crystalline Na_2HPO_4 external; (d) lyophilized Na_2HPO_4 external; (e) 50 μL ^{13}C-iodomethane.

Compared to ^{13}C-methylated ribonuclease A using 25 μL ^{13}C-iodomethane (Figure 3-15 (a)), there was a slight increase in the extent of methylation of ribonuclease A reacted in the presence (external) of NaOH or crystalline Na_2HPO_4 (6%); however, a more significant increase (approximately 20%) was observed in the presence of Na_2HPO_4 powder that was lyophilized (1 mL 0.2M Na_2HPO_4, pH 9.5; i.e., LpH 9.5), but a slight decrease in the ^{13}C-trimethylamino resonance of the lysyl ε-amino groups was evidenced (cf. Figure 3-15 (a), (c), and (d)). The latter reaction conditions revealed an extent of methylation that approached that of ribonuclease A reacted *in vacuo* with 50 μL ^{13}C-iodomethane (cf. Figure 3-15 (d) and (b)).

3.4.1.5 Effect of the Counterion on the Extent of *In Vacuo* ^{13}C-Methylation Reaction of Unbuffered Ribonuclease A

In an attempt to promote the reaction of carboxylate salt groups in protein with ^{13}C-iodomethane, the effect of the counterion of the carboxylate salt on the methylation reaction of ribonuclease A, LpH 6 without buffer was studied. Cesium carboxylate salts have been shown to react in an S_N2 alkylation to a greater extent than lithium or sodium carboxylate salts (March, 1985 (d), Kruizinga *et al.*, 1981). A carboxylate anion of a cesium carboxylate salt in a protein molecule (i.e., Cs^+ $^-$OOC—protein) has more ionic character than a carboxylate anion of a lithium or sodium carboxylate salt. Lithium or sodium are smaller atoms than cesium and form a covalent-like bond with a carboxylate anion (i.e., $Li^{\delta+}$- $^{\delta-}$OOC—protein). The carboxylate anion in a cesium carboxylate salt has more of a lone or bare anionic character compared to that in a lithium or sodium carboxylate salt. Hence, the cesium carboxylate salt would better facilitate the carboxylate anion to attack the electrophilic reagent iodomethane. Similarly, a bulkier

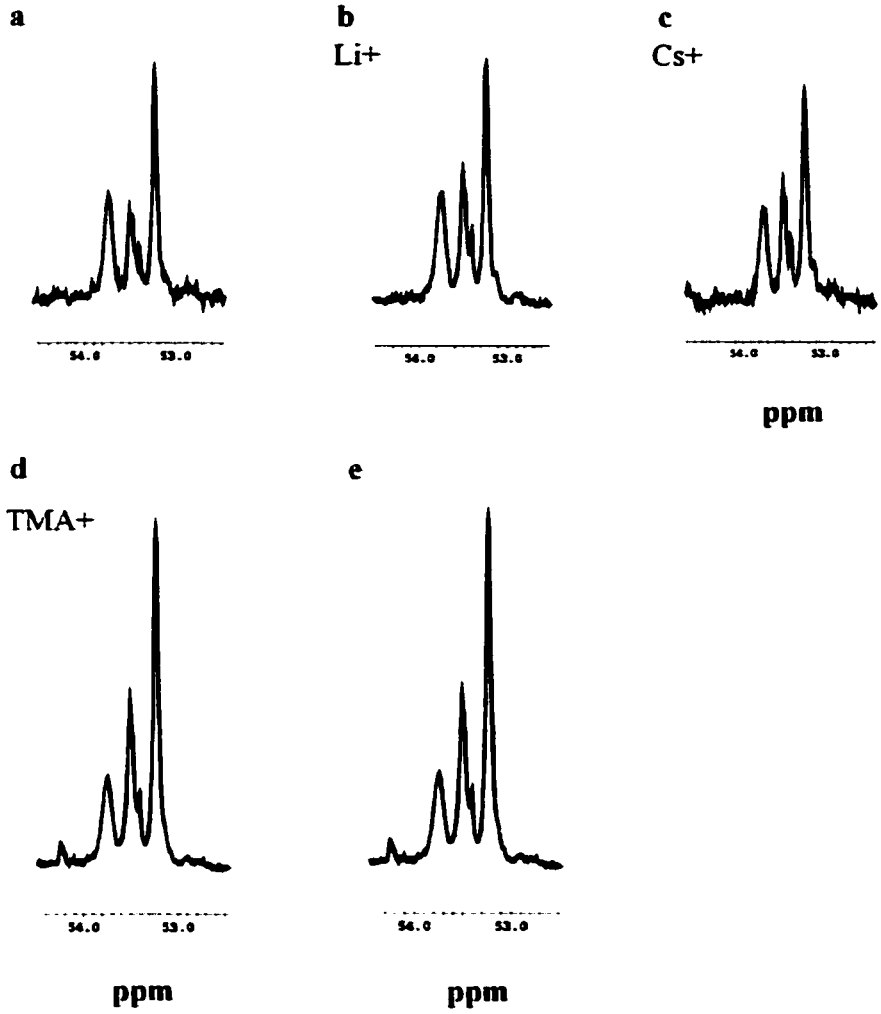


Figure 3-16. Effect of the counterion on the extent of *in vacuo* ^{13}C-methylation of unbuffered ribonuclease A (20 mg) at LpH 6.

Typical reaction conditions: 25 μL ^{13}C-iodomethane, 75°C, 24 h. (a) control, Na^+; (b) Li^+; (c) Cs^+; (d) TMA^+; cf. sample reacted with 50 μL ^{13}C-iodomethane (e).

counterion such as tetramethylammonium cation (TMA^+) would form a bond with more ionic character with a carboxylate anion cf. lithium and sodium, and is expected to promote the methylation of protein carboxyl groups. Counterions tested were Li^+, Cs^+ and TMA^+ (cf. control sample, which was assumed to contain mainly Na^+ counterions). ^{13}C-Methylated ribonuclease A samples prepared with these counterions were compared to ^{13}C-methylated ribonuclease A, LpH 6 prepared without buffer (control). Samples containing Li^+, Cs^+, or TMA^+ counterions for protein carboxylate anions were observed to have similar reactivity toward ^{13}C-iodomethane, but a higher reactivity compared to that of the control sample (cf. Figure 3-16 (a)-(d) γ-carboxyl ^{13}C-methyl ester resonances at 53.4-53.5 ppm); the samples containing Li^+ or Cs^+ counterions were also observed to have slightly higher extents of reaction for N^{ε}- or N^{α}-trimethyl amino groups of lysyl residues. The resonance of the latter group increased in peak intensity, which may be due to an increase in reactivity of δ-carboxyl groups toward ^{13}C-iodomethane, as the resonances of their ^{13}C-methyl ester derivatives overlap those of the N-terminal N^{α}-trimethyl amino groups of lysyl residues. For the ribonuclease A sample containing TMA^+ as counterion, the extent of reaction for N^{ε}-trimethyl amino groups of lysyl residues was slightly less than that for the ribonuclease A sample containing Li^+ as counterion or for that of the control. Perhaps the bulky nature of TMA^+ interferes with the ability of the N^{ε}- (or N^{α}-) amino groups to react with iodomethane; this may occur if the amino group normally resides close to a carboxylate anion, which now has TMA^+ as its counterion in the native protein structure. The amino group is likely to be in a different conformation and may not be exposed compared to the native protein lyophilized in the absence of TMA^+, which would lead to a decrease in reactivity of the

N^{ε}- or (N^{α}-) amino groups. It is noteworthy to compare the results of the counterion experiment with the ^{13}C-methylated ribonuclease A, LpH 6 (without buffer) sample, which was reacted with 50 μL ^{13}C-iodomethane. The intensity of the resonance for γ-carboxyl ^{13}C-methyl esters in samples with Li^{+}, Cs^{+}, or TMA^{+} counterions approached that for the sample reacted with 50 μL reagent; however, the extent of reaction of carboxylate groups was still less than the latter sample (cf. Figure 3-16 (a)-(d) with (e)).

3.4.1.6 Effect of Consecutive *In Vacuo* Reactions (^{12}C/^{13}C or ^{13}C/^{13}C) on the Extent of *In Vacuo* Methylation of Unbuffered Ribonuclease A

To test the hypothesis that liberated hydroiodic acid inhibits the *in vacuo* methylation reaction in unbuffered lyophilized protein samples, consecutive *in vacuo* methylation reactions were conducted on ribonuclease A. ^{12}C-Iodomethane and ^{13}C-iodomethane were used in the first and second *in vacuo* methylation reactions, respectively, for the control sample. The same was done for a second sample except that the sample was re-lyophilized at LpH 6 in between *in vacuo* methylation reactions. Re-lyophilization in between *in vacuo* methylation reactions served two purposes. One was to re-establish the LpH of the protein sample to the original value, and the other was to expose the reagent to a proportion of the reactive functional groups that may have been buried originally in the protein lyophilized powder, and hence, inaccessible to reagent. The results are shown in Figure 3-17. The control sample showed very little additional reaction with ^{13}C-iodomethane (2-3%) while significant (approximately 52%) additional reaction was observed for the sample that was re-lyophilized in between *in vacuo* methylation reactions. ^{13}C-NMR spectra for both consecutive reaction samples (2x 24 h) were compared to that for ribonuclease A reacted for 24 h with ^{13}C-iodomethane. Hence,

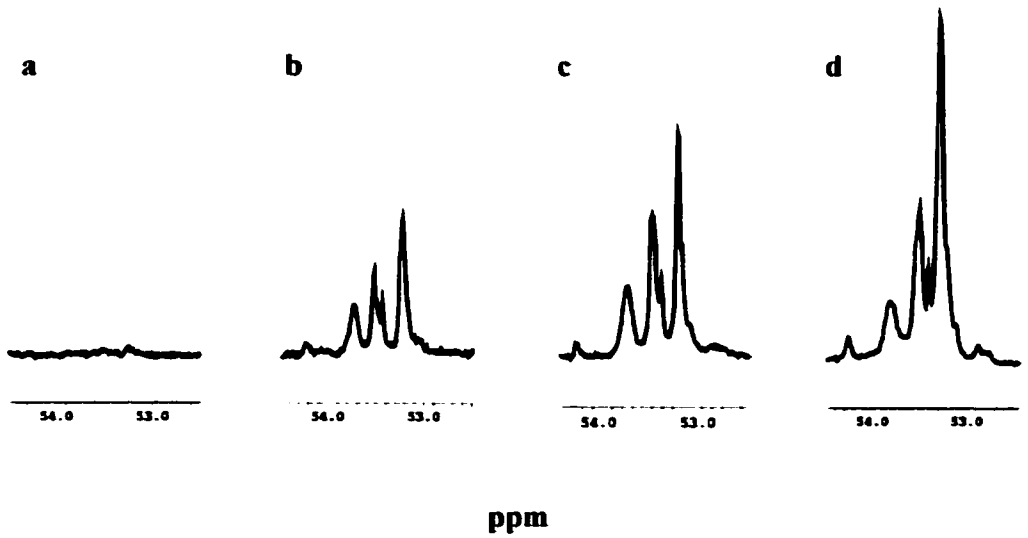


Figure 3-17. Effect of surface accessibility experiment on the extent of *in vacuo* consecutive $^{12}C/^{13}C$-methylation of unbuffered ribonuclease A (20 mg) at LpH 6. (a) Control reaction conditions for consecutive reaction: first *in vacuo* reaction (25 μL ^{12}C-iodomethane, 24 h, 75°C), followed by a second consecutive *in vacuo* reaction (25 μL ^{13}C-iodomethane, 2x24 h total, 75°C). (b) Sample reaction conditions: first *in vacuo* reaction (25 μL ^{12}C-iodomethane, 24 h, 75°C), followed by re-lyophilization at LpH 6, and second consecutive *in vacuo* reaction (25 μL ^{13}C-iodomethane, 2x24 h total, 75°C). Compare with (c) control sample for a typical ^{13}C-methylation reaction (25 μL ^{13}C-iodomethane, 24 h, 75°C) and (d) consecutive reaction with $^{13}C/^{13}C$-iodomethane (with re-lyophilization at LpH 6 in between reactions; 2x24 h total, 75°C).

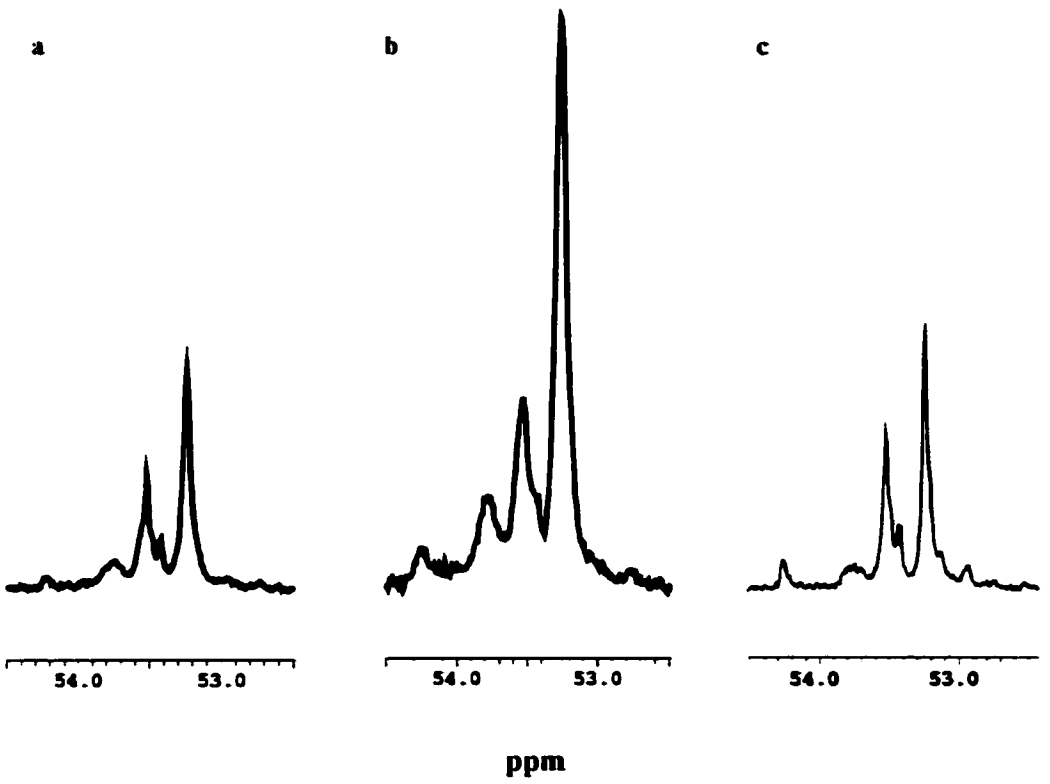


Figure 3-18. Effect of surface accessibility experiment on the extent of *in vacuo* consecutive $^{13}C/^{13}C$-methylation of unbuffered ribonuclease A (20 mg) at LpH 6.
(a) Control reaction conditions for consecutive reaction: first *in vacuo* reaction (25 μL ^{13}C-iodomethane, 24 h, 75°C), followed by a second consecutive *in vacuo* reaction (25 μL ^{13}C-iodomethane, 2x24 h total, 75°C). (b) Sample reaction conditions: first *in vacuo* reaction (25 μL ^{13}C-iodomethane, 24 h, 75°C), followed by re-lyophilization at LpH 6, and second consecutive *in vacuo* reaction (25 μL ^{13}C-iodomethane, 2x24 h total, 75°C). Compare with (c) control sample for a typical ^{13}C-methylation reaction (25 μL ^{13}C-iodomethane, 24 h, 75°C).

the extent of derivatization is directly related to the surface accessibility of potentially reactive functional groups of the protein molecules in the lyophilized powder, as well as to the restoration of the sample LpH.

Alternatively, ^{13}C-iodomethane was used for both first and second *in vacuo* methylation reactions. The results are shown in Figure 3-18. There was very little difference in the extent of reaction for the 24 h control sample and the 2x24 h consecutive reaction control sample of ^{13}C-methylated ribonuclease A. The 2x24 h ^{13}C-methylation reaction with re-lyophilization in between *in vacuo* reactions (at LpH 6) showed a significant increase in the extent of methylation (53%).

3.4.1.7 Estimation of the Extent of ^{13}C-Methylation of Carboxylate Groups in Ribonuclease A by High-Voltage Paper Electrophoresis

The extent of methylation of carboxylate groups in ribonuclease A, methylated at LpH 3, 6 and 10 with ^{13}C-iodomethane, was estimated by high-voltage paper electrophoresis from the electrophoretic mobility at pH 6.5 of peptic peptides of digested derivatized protein. At pH 6.5, C-terminal peptides with α-, γ-, or δ-carboxyl groups of Asp and Glu residues fully methylated will possess an overall positive charge, and will migrate on paper as basic peptides. The same will be true for peptides containing no Asp or Glu residues (having a methylated C-terminal residue). Non-C-terminal peptides having fully methylated γ- or δ-carboxyl groups will be overall neutral or positive in charge, and will migrate in the neutral or basic band on paper. Non-C-terminal peptides containing Asp or Glu residues that are not methylated may have an overall negative charge, and will migrate as acidic peptides on paper (but could still be neutral or positive). Hence, a pepsin digest of fully methylated esterified protein should not contain acidic peptides. If there is any methylation, the intensity of the acidic peptide band

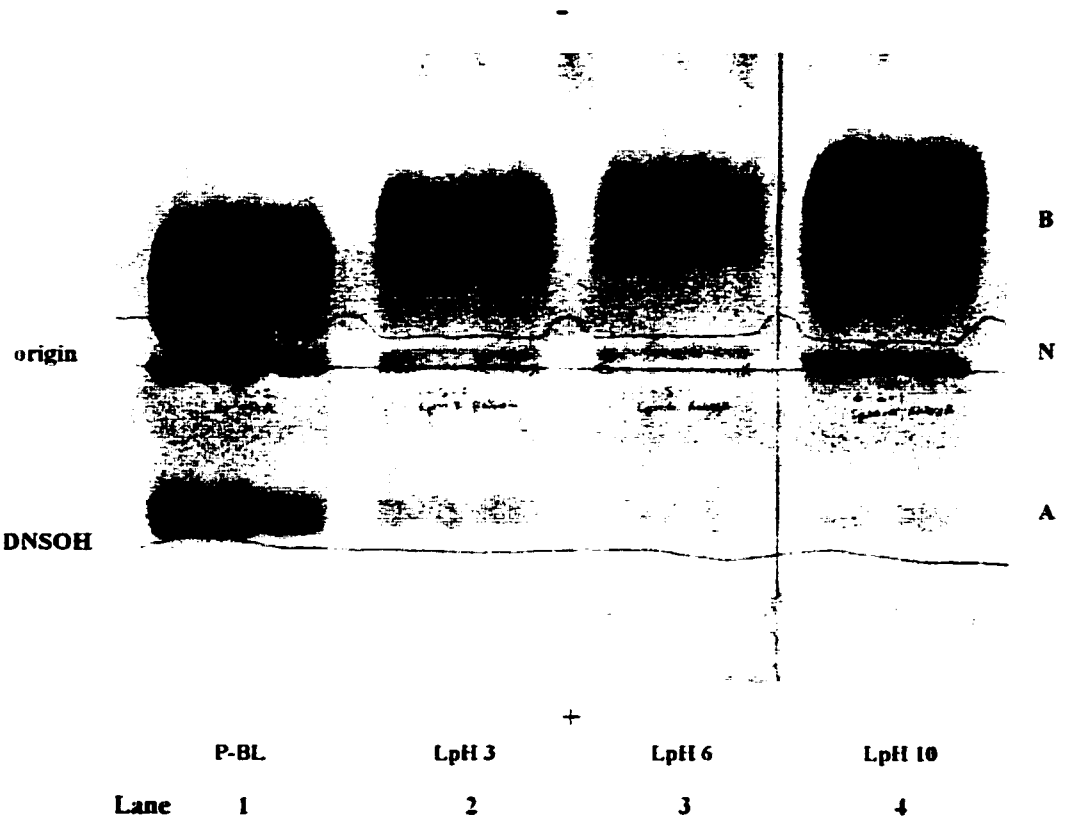


Figure 3-19. High-voltage paper electrophoresis of peptic peptides of digested methylated ribonuclease A at pH 6.5 for the estimation of the extent of methylation of carboxylate groups.

Lane 1, unmethylated ribonuclease A control (P-BL = protein blank/unmethylated); lanes 2, 3, 4, ribonuclease A methylated at LpH 3, 6, and 10, respectively.

B, N, A = basic, neutral, acidic bands; DNSOH = dansyl sulfonic acid. Bands were detected with ninhydrin/cadmium acetate/acetone reagent.

Table 3-5. Extent of Modification of Carboxyl Groups of Esterified Ribonuclease A[†]

Ribonuclease A	Reagent(s)	Extent of Modification[a]
LpH	^{13}C-Iodomethane (µL)	Extent of Methylation %
3	25	14.2 ± 7.3[b]
4	25	16.0 ± 11.3
5	25	27.7 ± 5.8
6	25	25.3 ± 10.2
7	25	37.4 ± 7.6
10	25	57.8 ± 10.1
	^{13}C-methanol/4 N HCl in dioxane (µL:µL)	Extent of Esterification %
3	25:10	59.7 ± 7.3[b]
6	25:5	12.0 ± 5.7
6	25:10	18.0 ± 2.7
6	25:20	41.7 ± 7.6
	^{13}C-ethanol/4 N HCl in dioxane (µL:µL)	
3	30:10	24.2 ± 7.2

[a] Calculated from the difference of titratable groups titrated from pH 3.0-5.5. Forward and backward titrations were carried out; the values given for the extent of methylation are based mainly on the backward titration data, whereas those for the extent of esterification are based on both forward and backward titration data. Percentages were calculated as follows:

$$\frac{[\#\mu L\ \text{P-BL (+ W-BL)}] - [\#\mu L\ \text{S (+ W-BL)}]}{[\#\mu L\ \text{P-BL (+ W-BL)}] - [\#\mu L\ \text{W-BL}]} \times 100\% = \frac{\text{P-BL} - \text{S}}{\text{P-BL}} \times 100\%$$

where P-BL=native protein blank, (+W-BL)=water blank inclusive in the P-BL or S, S=sample, W-BL=water blank as determined in the absence of P-BL or S.
Data was normalized with respect to mass of protein used.

[b] Standard deviation = $s = \sqrt{\Sigma_i (x_i - x')^2/(n-1)}$, where x_i is the individual measured values, x' is the arithmetic mean, and n is the number of measured values.

[†] Raw data can be found in Appendix 3-1, Table 3-5A.

should be less than that of the control, peptic peptides derived from digested unmethylated protein, and its intensity should decrease with the extent of reacted carboxyl groups (i.e., darker band, less reacted; lighter band, more reacted). The results obtained (Figure 3-19) revealed that methylation was incomplete at LpH 3, 6 and 10, but there was a progressive increase in extent of methylation with LpH, which supported the results obtained from the difference titration method (Table 3-5).

3.4.2 Extent of Esterification of Ribonuclease A Carboxylic Acid Groups with ^{13}C-Methanol/HCl

The ^{13}C-NMR spectrum of ^{13}C-methyl-esterified ribonuclease A (LpH 3; ^{13}C-methanol/ 4 N HCl in dioxane 25μL:10μL) is shown in Figure 3-8a. Comparison with the spectrum of ^{13}C-methylated ribonuclease A (25μL ^{13}C-iodomethane) revealed a larger (88% greater) extent of reaction of carboxyl groups for the *in vacuo* esterification reaction with ^{13}C-methanol/HCl than that for the *in vacuo* methylation reaction with ^{13}C-iodomethane (cf. Figure 3-14 (a)). Some variability in the extent of esterification with ^{13}C-methanol/HCl at LpH 3 was observed; however, the latter reaction always revealed (semi-quantitatively by ^{13}C-NMR spectroscopy) a larger extent of reaction (by 88%) cf. the *in vacuo* methylation reaction with ^{13}C-iodomethane. As mentioned above, one reason for the variability in the extent of reaction could be the amount of water present in the lyophilized powder. Water shifts the equilibrium of the esterification reaction toward hydrolysis (see Figure 3-12). In reactions with alcohol/HCl, the extent of esterification was not greatly affected by the LpH of the protein.

3.4.2.1 Effect of Acid Catalyst Concentration on the Extent of *In Vacuo* ^{13}C-Methyl-Esterification of Ribonuclease A Carboxylic Acid Groups

The effect of the amount of acid catalyst on the *in vacuo* esterification reaction with ribonuclease A at LpH 3 with ^{13}C-methanol was examined. A small degree of reaction with ^{13}C-methanol was evidenced with ribonuclease A at LpH 3 in the absence of the acid catalyst; hence, sufficient acid was present in the lyophilized protein to effect partial esterification. No esterification was observed with ribonuclease A at LpH 6 in the absence of acid catalyst. The largest extent of reaction at LpH 3 was observed with 25μL:10μL ^{13}C-methanol/4 N HCl in dioxane (see Figure 3-20). A higher amount of acid (20 μL) resulted in additional resonances in the α-carboxyl ^{13}C-methyl ester region of the ^{13}C-NMR spectrum. These extra α-carboxyl ^{13}C-methyl ester resonances were observed also for longer reaction times and for consecutive reactions (see below). They are likely to originate from the acid-catalyzed hydrolysis (March, 1985 (b); AAC2 mechanism; protonation of the amide oxygen atom in an acidic environment renders the carbonyl carbon more susceptible to nucleophilic attack by the little water present in the *in vacuo* reaction conditions) of the peptide backbone and the subsequent esterification of the newly formed carboxylic acid groups.

3.4.2.2 Estimation of Extent of *In Vacuo* ^{13}C-Methyl-Esterification of Ribonuclease A Carboxylic Acid Groups by Difference Titration

The extent of esterification of ribonuclease A using ^{13}C-methanol and acid catalyst as determined by the difference titrimetric method of the esterified and nonesterified protein (titrated between pH 3.0 and 5.5) was calculated to be 59.7±7.3% (for a single *in vacuo* reaction (24 h, 75°C; Table 3-5). Reaction with ^{13}C-ethanol/HCl resulted in 25% esterification of carboxyl groups under the same conditions. The ^{13}C-NMR spectra of half of each esterified sample used for the estimation of the extent of

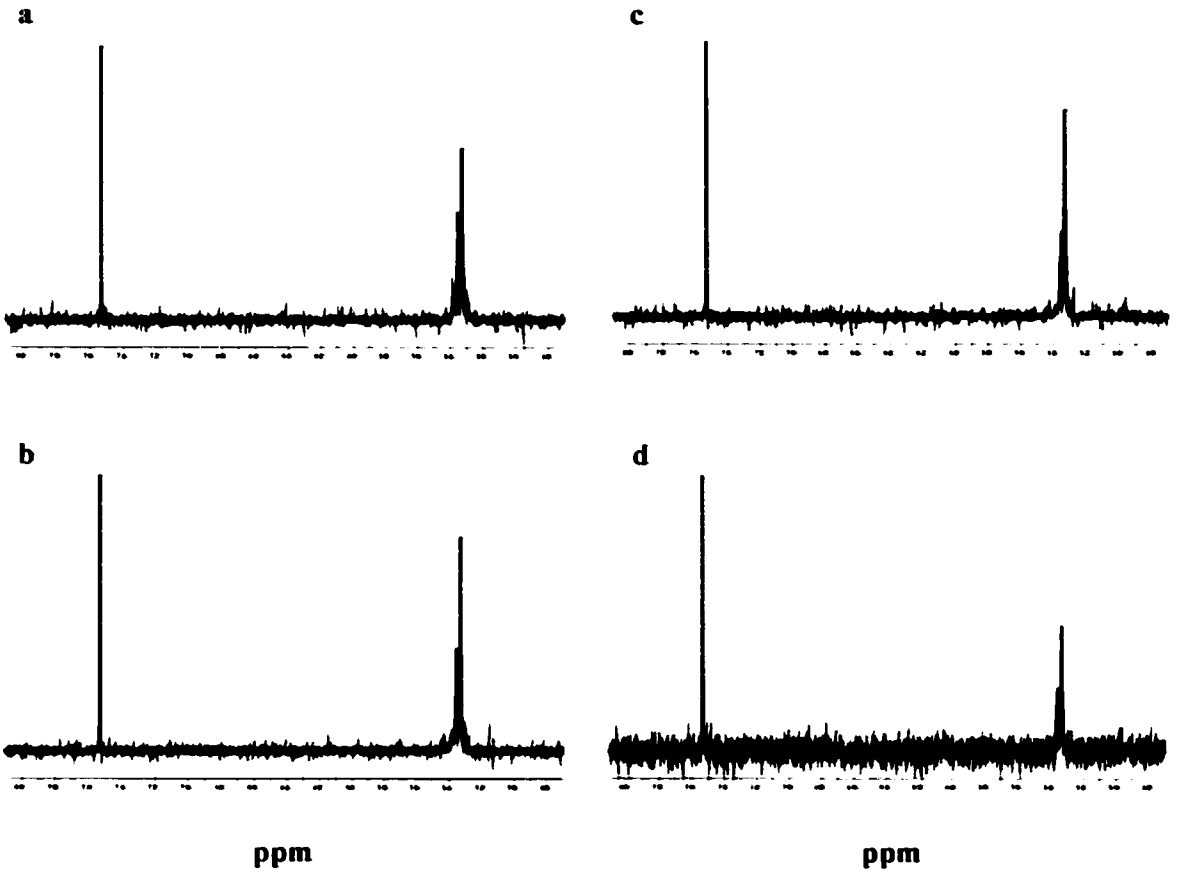


Figure 3-20. Effect of acid catalyst concentration on the extent of *in vacuo* ^{13}C-methyl-esterification of ribonuclease A (20 mg) at LpH 3.

(a) 2.5 μL; (b) 5 μL; (c) 10 μL; (d) 20 μL added 4 N HCl in dioxane.

Typical reaction conditions: 25 μL ^{13}C-methanol, 24 h, 75°C.

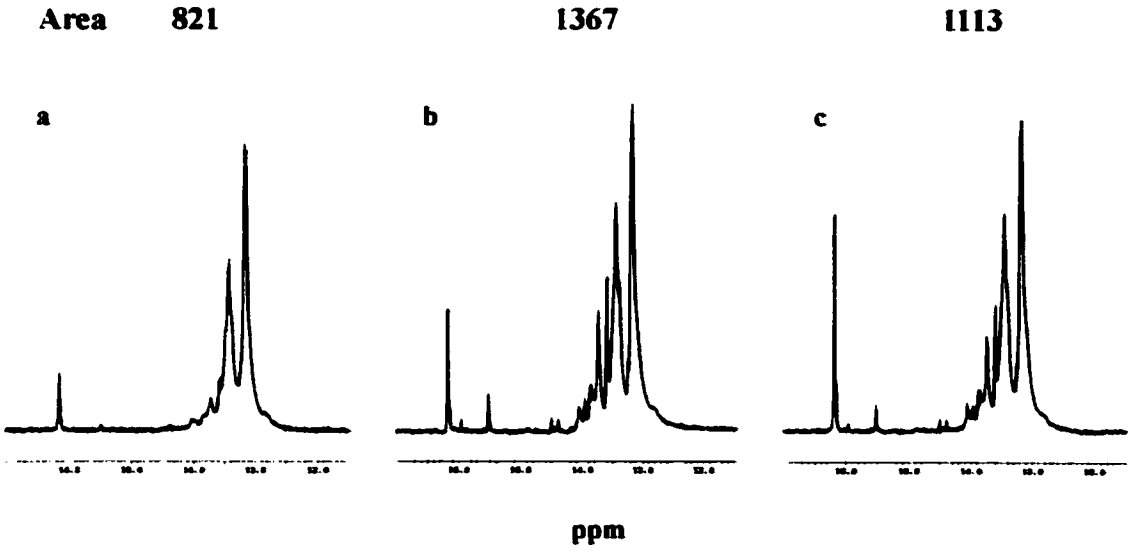


Figure 3-21. Effect of surface accessibility experiment on the extent of *in vacuo* consecutive $^{13}C/^{13}C$-methyl-esterification of unbuffered ribonuclease A (20 mg) at LpH 3 with ^{13}C-methanol/HCl.

(a) Non-consecutive reaction control; (b) consecutive reaction without freeze-drying in between reactions (2x24 h total, 75°C); (c) consecutive reaction with re-lyophilization at LpH 3 in between reactions (2x24 h total, 75°C).

esterification by difference titration revealed that there is variability in the extent of reaction from sample to sample. Consecutive reaction experiments with ^{13}C-methyl-esterified ribonuclease A also revealed variable results for the extent of reaction by comparison of ^{13}C-methyl ester peak areas (see Sections 2.5.2 and 2.12; Table 3-6). The main reason for a decrease in the extent of reaction is believed to be the amount of water trapped in the lyophilized powder, which is related to the time of lyophilization (see Section 3.4.1).

The results of the consecutive esterification reaction experiments are shown in Figure 3-21 and Table 3-6. The apparent maximum achievable extent of esterification was observed for the consecutive reactions without re-lyophilizing the sample at LpH 3 in between *in vacuo* reactions. A significant increase in the extent of esterification of carboxyl groups was observed for consecutive reactions. In contrast to the reaction of the carboxyl groups with iodomethane, redissolving and re-lyophilizing the esterified protein did not result in a significant increase in the extent of methylation of carboxyl groups on subsequent reaction with methanol/HCl. Interestingly, the maximum value of 1362 and 1388 area units was reached for 10 mg and 20 mg samples of ribonuclease A (with CR1-C2 samples), respectively, (see Table 3-6 legend); both samples originally were reacted as 30 mg and subsequently split into 3x 10 mg portions (NCR-C1, CR1-C2, CR2) or 1x 10 mg (CR1-C2) and 1x 20 mg (CR1-C2) portions, respectively. These results, plus the fact that the area for the 20 mg NCR-C1 sample did not correspond to twice that for the 10 mg NCR-C1 sample suggests that the *in vacuo* esterification reaction is inhomogeneous. Re-lyophilizing the sample at LpH 3 in between *in vacuo* reactions resulted in a slight decrease in the extent of esterification (see CR2 areas, Table 3-6).

Table 3-6. Consecutive *In Vacuo* ^{13}C-Methyl-Esterification Reaction Experiments

Ribonuclease A (ND or D)[a]; Initial Reaction Mass (mg)	Mass (mg)	NCR-C1[b] Area	CR1-C2[c] Area	CR2[d] Area
ND;20	10	-	272	384
D;20	10	-		
☼ND;30	10	1237	1362	1173
ND;30	10	558	1042	868
D;30	10	634	890	708
☼D;30	10	535	1216	1226
ND;20	20	374	367	610
ND;30	10	A 506		
	10	B 653		
	10	C 468		
☼	20	A 734	B 1388	C 1170
D;30	10	A 664		
	10	B 483		
	10	C 442		
☼	20	A 821	B 1367	C 1113

[a] ND = not dialyzed; D = dialyzed

[b] NCR-C1 = no consecutive reaction - control 1; areas for samples compared to area of 100 for internal standard

[c] CR1-C2 = consecutive reaction 1 - control 2

[d] CR2 = consecutive reaction 2

☼ best results observed; protein samples were lyophilized extensively to dryness to ensure very little water trapped in the lyophilized protein cake; less than optimal results were due to poorly dried lyophilized protein samples; the variable results clearly indicate the importance of limiting the water content in the lyophilized protein. Use of dialyzed or non-dialyzed protein was not a factor in the extent of reaction.

This was probably the result of the re-addition of water to the derivatized protein; trapped water molecules in the re-lyophilized protein matrix were likely present in a slightly greater extent compared to that in the CR1-C2 sample (without re-lyophilization in between *in vacuo* reactions). Dialysis of the protein sample prior to reaction does not seem to be a factor with respect to the variability in the results (see Table 3-6). How the protein molecules pack in the lyophilized cake seems to play a key role in the extent of modification. Hence, further reactions on dispersed protein powder (as opposed to cake) and studies on lyophilization volume are required to observe the optimal extent of esterification.

3.5 Determination of the Structural Integrity of Esterified Ribonuclease A via SDS-PAGE

Figure 3-22 shows the SDS-PAGE PhastGel™ gels of ribonuclease A following *in vacuo* esterification with methanol or ethanol and hydrochloric acid catalyst. The gas phase esterification reaction did not have a detrimental affect on the protein structure and no breakdown of the protein had occurred, as the mobility, and hence the molecular weight, of both the native and esterified proteins were the same. However, a small amount of higher molecular weight material was observed, which suggested that a small amount of crosslinking may have occurred. Figure 3-22(a) (lanes 4, 5) shows that some insoluble material did not enter the gel for the methyl-esterified ribonuclease A samples; this was not the case with a gel run subsequently to the one in question (gel not shown) using the same samples. Hence, since the samples were not boiled prior to loading to preserve the methyl or ethyl esters of protein carboxyl groups, it was required to wait a few minutes (e.g., 30 min) at room temperature for the protein molecules to be sufficiently solubilized by SDS. The SDS-PAGE gel shown in Figure 3-22(b) was

prepared with a high loading of both native and methyl-esterified ribonuclease A, and more clearly illustrates the higher molecular weight material, which may originate from intermolecular crosslinking by a nucleophilic unprotonated amino group (lysyl residues) displacing the methoxyl group of a carboxyl methyl ester. Although the number of unprotonated amino groups is small at low LpH, evidence provided from the *in vacuo* methylation of human albumin with iodomethane for their presence in the lyophilized protein structure has been observed in the present study (see Figure 3-3 human albumin, LpH 3, 4 or 5).

The mobility of the methyl-esterified protein (Figure 3-22 (b)) is slightly greater than that of the native ribonuclease A, which is likely due to the more overall positive nature of the derivatized protein compared to the native protein. The derivatized protein can bind relatively more SDS than the native protein and thus migrate further in the gel. Such an occasional artifact has been shown to be possible in the widely used analytical method of SDS-PAGE (McCumber and Clem, 1976). Some of the other causes for anomalous electrophoretic behaviour are discussed in Chapter 6, Part B (Vakos, Burton and Kaplan, 1997). Intramolecularly crosslinked protein may also yield a falsely low molecular weight because it would not denature properly in SDS (Steele and Nielson, 1978).

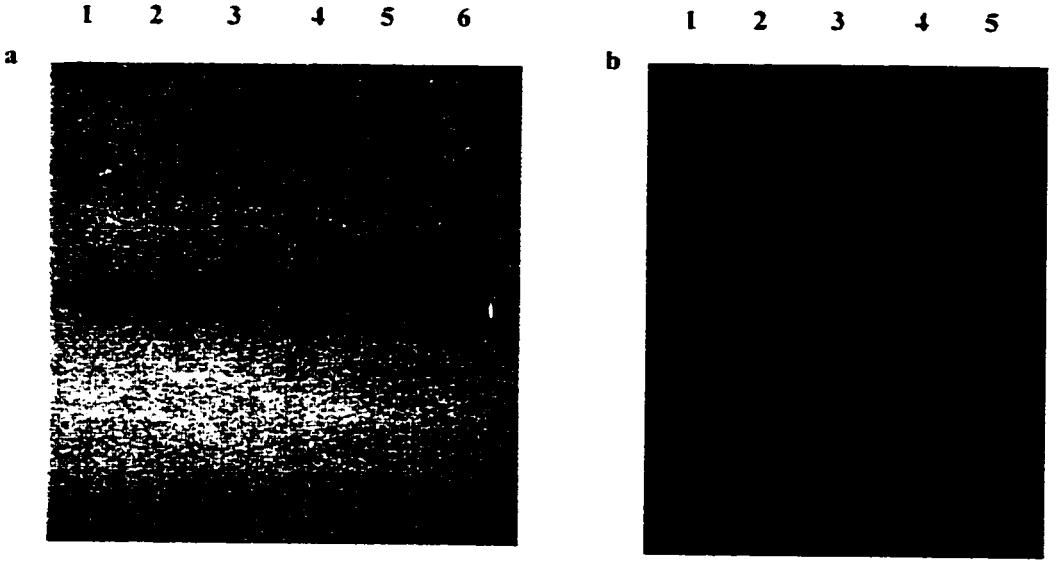


Figure 3-22 ***In Vacuo*** **Methyl- and Ethyl-Esterification of Ribonuclease A Carboxyl Groups.**

(a) Samples (2 μg) were not heated and were run on an SDS-PAGE 20% homogeneous PhastGel™ gel. Lanes 1, 2, 6: native ribonuclease A; Lane 3: ethyl-esterified ribonuclease A; Lanes 4, 5: methyl-esterified ribonuclease A.

(b) Samples (4 μg) were not heated and were run on an SDS-PAGE 20% homogeneous PhastGel™ gel. Lanes 1, 2, 5: native ribonuclease A; Lanes 3, 4: methyl-esterified ribonuclease A.

4. Discussion

The reagents used in this study to effect methyl or ethyl ester derivatives of protein carboxyl groups, namely iodomethane and anhydrous methanol or ethanol with hydrochloric acid (HCl) catalyst, have been used previously to a limited extent to chemically modify proteins (Blackburn, Hemphill Carter, and Phillips, 1941; Fraenkel-Conrat and Olcott, 1945; Ram and Maurer, 1959; Broomfield *et al.*, 1965; Link and Stark, 1968; Acharya and Vithayathil, 1975). The use of iodomethane to modify protein carboxyl groups has been restricted (Blackburn, Hemphill Carter, and Phillips, 1941); probably because this reagent effects modification of other functional groups and the carboxyl ester derivatives are unstable under the conditions normally used (i.e., slightly alkaline conditions). The use of methanol and hydrochloric acid catalyst to convert protein carboxyl groups to methyl esters results in few chemical changes other than the esterification.

The results of the present study demonstrate that carboxyl groups in lyophilized proteins can be esterified *in vacuo* with the alcohol and the HCl catalyst in the vapour phase. With methanol/HCl approximately 60-70% esterification was achieved. This is very similar to that achieved by suspension of the protein in acidic methanol (Broomfield *et al.*, 1965). However, the *in vacuo* procedure has several advantages over the conventional esterification protocol: 1) The esterification is much more rapid, as the reaction can be carried out at elevated temperatures, under which conditions lyophilized proteins are stable (Zaks and Klibanov, 1984, 1985; Ayala *et al.*, 1986; Reslow *et al.*, 1987; Mozhaev *et al.*, 1991; Volkin *et al.*, 1991). The conventional esterification procedure takes anywhere from one day to one week depending on conditions (Fraenkel-

Conrat and Olcott, 1945; Ram and Maurer, 1959; Broomfield *et al.*, 1965; Acharya and Vithayathil, 1975), whereas the *in vacuo* procedure takes 2 to 4 h at 75°C. 2) Very small amounts of alcohol are required, approximately 1 μL per mg of protein. This makes it feasible to use labeled alcohols to incorporate ^{14}C, ^{13}C, ^{2}H and ^{3}H labels into the esterified protein in a cost-effective manner. 3) The procedure can be applied to a very small amount of protein, as the protein is not suspended in alcohol and remains in the solid state throughout the procedure. 4) Any volatile alcohol can be used in the esterification procedure. 5) The likelihood of protein denaturation or breakdown is minimized, as the protein remains in the lyophilized state throughout the procedure.

^{13}C-Labeled methanol and ethanol were used to incorporate an NMR active nucleus into the protein. It was found that the chemical shifts of the resonances of the α-, γ-, and δ-carboxyl methyl esters were sufficiently different that each of these classes of carboxyl ester in the protein could be unequivocally identified. The resonances of the γ- and δ-carboxyl esters in the protein usually appeared as a single peak, only slightly broadened, indicating that the chemical shifts of these groups are not very sensitive to their environments within the protein. In contrast, the resonances of the individual esterified α-carboxyl groups showed a strong dependence on the nature of the C-terminal amino acid and sequence (i.e., penultimate amino acid residue). In the case of insulin, it was possible to differentiate two forms of this protein that differed only in the penultimate amino acid of the B-chain by the difference in the chemical shifts of the C-terminal carboxyl ^{13}C-methyl esters. Incorporation of such a ^{13}C-label into α-carboxyl groups provides a new approach for determining if a pure protein consists of more than one polypeptide chain, or if a purified protein contains a significant amount of a

contaminating protein of similar size, which would be difficult to detect by SDS-PAGE. Although we have not as yet determined the chemical shifts expected from all 20 C-termini, it appears that it should be possible to make a tentative identification of the C-terminus based on such chemical shifts. Until such data is available, evidence for the presence of a particular amino acid at the C-terminus of a protein can be obtained by comparing the chemical shift of its ^{13}C-methyl ester with that of an esterified acetylated dipeptide having the appropriate C-terminus.

Ethanol/HCl also esterified carboxyl groups in the lyophilized protein but to a lesser extent (Figs. 2b and 3b) than methanol/HCl. It appears that the bulkiness of the alcohol is a factor in determining the extent of esterification. Nevertheless, the extent of esterification with ethanol labeled with carbon-13 in the C_1 position was sufficient for NMR spectroscopic analysis. As in the case of the methyl esters, the chemical shifts of the resonances of the ethyl esters showed a similar pattern for the three classes of carboxyl groups in the protein. *In vacuo* esterification with ethanol/HCl could be very useful in cases where it is desirable to form a more stable ester derivative, or to introduce a more stable ester derivative containing a ^{14}C, ^{13}C, CD_2, CD_3, CT_2 or CT_3 label into the protein.

Iodomethane was shown to react *in vacuo* with deprotonated carboxyl groups in lyophilized proteins to form methyl esters with all three classes of carboxyl groups. In contrast with *in vacuo* esterification with gaseous acidic alcohol, the extent of methylation of the carboxyl groups with iodomethane showed a strong dependence on the pH of lyophilization (LpH) and confirmed the pH memory effect. In Chapter 2, it was shown that the amino groups in proteins are readily trimethylated by *in vacuo* reaction

with iodomethane and as a result it is difficult to achieve selective methylation (esterification) of carboxyl groups. Nevertheless, at LpH values between 3 and 4, the reaction is preferential for the carboxyl groups in proteins. If extensive selective esterification of carboxyl groups is desired, then *in vacuo* reaction with methanol/HCl is the method of choice. However, it is much easier to control the extent of reaction with iodomethane so in cases where a low extent of esterification of carboxyl groups is desired, reaction with iodomethane would be preferable.

The suspension of protein in acidic methanol usually results in a considerable change in conformation, denaturation, and inactivation (Broomfield *et al.*, 1965; Means and Feeney, 1971; Acharya and Vithayathil, 1975). Charge interactions in proteins are in part responsible for the maintenance of structure and interaction energies depend on the dielectric strength of the medium (Creighton, 1993; Fersht, 1985). The *in vacuo* methyl-esterification reaction of lyophilized protein with methanol and hydrochloric acid catalyst in the vapour phase results in the introduction of a small group into the carboxylic acid function. Depending on the extent of reaction, the chemical modification could result in a substantial removal of surface negative charge, which may alter protein conformation about the active site (when the derivatized protein is placed in aqueous medium) or affect activity by methylation of one or more catalytically important carboxyl group(s). If the extent of esterification is large (e.g., carboxyl groups of ribonuclease A were modified 60-70% via *in vacuo* esterification with ^{13}C-methanol/HCl, which was comparable to that achieved by suspension of the protein in acidic methanol (Broomfield *et al.*, 1965)), then a large change in the local and global electrostatic character of the protein will result. The protein molecule will become more hydrophobic at neutral pH,

and overall positively charged at low pH values when amino groups are protonated. When the extent of esterification is controlled and limited (e.g., via *in vacuo* esterification with ^{13}C-iodomethane at low LpH values or with ^{13}C-methanol/HCl for short incubation times), the original charge may not be drastically perturbed and the modified protein should be native-like and partially active (e.g., Acharya and Vithayathil, 1975). Thus, enzyme activity and structural studies are necessary on various esterified protein derivatives to verify structural integrity and the maintenance of catalytic activity.

Broomfield *et al.* (1965) have shown that eight out of eleven (approximately 73%) carboxyl groups of ribonuclease A can be esterified by suspension of the protein in acidic methanol. The comparable extent of chemical modification by the *in vacuo* technique to that obtained by conventional esterification supports the observation that lyophilized proteins maintain at least in part their native-like structure, and suggests that the derivatized groups are "exposed" while the underivatized groups are "buried" in the hydrophobic core of the protein molecule (see Chapter 1). Other evidence stemming from the *in vacuo* chemical modification of lyophilized proteins that supports the idea that structural integrity is maintained is the observation that for several model proteins (except insulin), Tyr residues, which are usually buried, did not react with ^{13}C-iodomethane at LpH 10, indicating that the protein structure is largely intact. In insulin, Tyr residues were shown to react *in vacuo* with ^{13}C-iodomethane, indicating that this small protein is partially unfolded at LpH 10. The research of Broomfield *et al.* (1965) is a good example of how esterification of protein carboxyl groups could be used to elucidate protein structure-function information. The aim of their studies was to elucidate the carboxyl groups with abnormally low pK_a values by esterifying the

"exposed" carboxyl groups under conditions where the "buried" residues do not react. The product contained three free Asp carboxyl groups and possessed essentially no enzymatic activity, which could be largely regained by mild saponification. The folded structure of the derivative was shown (by difference spectrophotometry) to be similar to that of the native protein. The three free Asp carboxyl groups were located in the sequence, which provided information about the role of specific tyrosine-carboxylate interactions in the ribonuclease A molecule (Riehm *et al.*, 1965).

Acharya and Vithayathil (1975) have characterized the initial course of esterification of carboxyl groups of ribonuclease A in acidic methanol. Esterification of nearly five of the eleven free carboxyl groups in the protein resulted in almost complete inactivation of the enzyme. The inactivation was shown to be linear with the chemical modification of carboxyl groups during the initial stages of esterification. Five of the initial products of methyl ester derivatives of ribonuclease A were partially active (100-20% of that for ribonuclease A with up to three carboxyl groups modified). The esterification proceeded sequentially with some specificity. The initial product of esterification with reduced catalytic activity (65% of that for ribonuclease A) is a dimethyl ester with Glu-49 and Asp-53 modified, indicating that one (i.e., Glu-49) or both residues is/are important in maintaining the catalytically active structure of the protein. Subsequently, the esterification reaction appeared to proceed preferentially at the C-terminal region of the molecule. Comparison of the reactivities of carboxyl groups of ribonuclease A in acidic methanol to that known in aqueous solutions (with carbodiimides) suggested that the structure of ribonuclease A in the nonaqueous solvent resembles, at least in part, the structure in aqueous environment.

The results of the present study for the *in vacuo* (75°C) esterification (with ^{13}C-methanol/HCl) time course experiment with ribonuclease A (LpH 3) (Figure 3-3A) can be further analyzed given the data of Acharya and Vithayathil (1975) and Broomfield *et al.* (1965). At short reaction times up to 1 h there is very little reaction of the α-carboxyl groups and it appears from the peak intensities that the extent of esterification of δ-carboxyl groups is greater than that of the γ-carboxyl groups, given that ribonuclease A has five Asp and five Glu residues. However, after only 15 min reaction time, it is evident that at least two of each γ- and δ-carboxyl groups react, as evidenced by the slight variation in the chemical shifts of the individual esterified side-chains of Asp and Glu residues. After 2 h, the C-terminal α-carboxyl group is modified. The initial rate of incorporation of ^{13}C-label was linear, as evidenced by integrated peak areas, which likely will result in progressively inactive products (Acharya and Vithayathil, 1975). Moreover, the products of the *in vacuo* esterification reaction are likely a mixture of derivatives, since the protein is lyophilized and kinetically trapped in several conformations (Zaks and Klibanov, 1988); and hence, should possess different degrees of activities. These results support those obtained with ribonuclease A suspended in acidic methanol, which indicated that Glu-49 and Asp-53 are modified first, followed by the preferential reaction of the C-terminal residue (Acharya and Vithayathil, 1975), and that the three free carboxyl groups were Asp residues when eight of the eleven carboxyl groups were modified in an extensive reaction (Broomfield *et al.*, 1965).

An important point to note is that the ester formation is readily reversible by dissolving the esterified protein and incubating at neutral or slightly alkaline pH values. Previous studies have shown that in cases where esterification results in loss of enzymatic

activity, hydrolysis of the methyl esters regenerates the activity (Broomfield *et al.*, 1965; Acharya and Vithayathil, 1975). Stewart *et al.* (1997) showed that chymotrypsin retained its activity after *in vacuo* methylation with iodomethane in the presence of ligands that protect the active site histidine residue from being dimethylated on the imidazole ring, but at the time did not know that the carboxyl groups were also being esterified. However, under the alkaline conditions of the enzymatic assay, the esters were hydrolysed to yield chymotrypsin methylated only at imidazole and amino groups. Such hydrolysis is of particular relevance to the *in vacuo* trimethylation of amino groups with iodomethane (Chapter 2; Vakos *et al.*, 2000) as there is some esterification of carboxyl groups at the alkaline LpH values used in this reaction. The present study shows that this is not a problem as the methyl esters are easily removed.

The choice of extent of esterification will depend on the application for which the protein is intended. A high extent of modification of carboxyl groups is desired if the activity of the protein is not a concern. The results of the present study have led to the development of a working strategy for C-terminal sequence identification. The discovery that iodomethane reacts *in vacuo* to methylate protein carboxyl groups paved the way for the isolation and identification (sequence determination) of the C-terminal peptide(s) of *in vacuo* methylated proteins employing the diagonal two-dimensional high-voltage paper electrophoresis technique of Duggleby and Kaplan (1975), autoradiography, and tandem mass spectrometry. Taralp (1997, Doctoral Thesis) and Clairmont (1998, Doctoral Thesis) addressed this problem in earlier investigations; however, their methods were not very successful. The development of the present methodology has allowed for Clairmont (1998, Doctoral Thesis) to have moderate success and refinements in the

protocol by Nicolas Stewart (Ph.D. thesis to be submitted) have led to a working strategy for C-terminal sequence identification. In the present study, *in vacuo* esterification of protein with methanol/HCl gave a very high level of derivatization of carboxylic acid groups, which should make it easier to identify and isolate C-terminal diagonal peptides. Although *in vacuo* esterification with ethanol/HCl gave a lower extent of modification (as determined by difference titration of methyl- and ethyl-esterified ribonuclease A; see Table 3-5), this chemical modification may improve the peptide isolation procedure because an ethyl ester is more stable than a methyl ester.

Furthermore, since a high extent of esterification would lead to an overall positively charged protein molecule, it is possible to form hydroxamates of protein methyl esters using hydroxylamine, which would have the advantage of being more stable than the methyl ester group, and would maintain some degree of polarity on the protein surface. It may be possible to improve the solubility of proteins, which are normally insoluble at neutral pH (e.g., insulin) by carrying out sequential *in vacuo* reactions with 1) iodomethane, and 2) methanol/acid, followed by 3) an aqueous hydroxylamine reaction. The derivatized protein would have permanent positively charged amino groups at all pH values, and would have polar hydroxamate groups formed from carboxyl methyl ester groups that would help to maintain the protein structure. An application of such a protein derivative with improved solubility may be to increase the therapeutic value of a pharmaceutical protein. Insulin is a good test protein for such a study because it is insoluble at neutral pH and several excipients must be added to the pharmaceutical preparation to enhance its solubility, as this hormone must be administered with a needle to be effective toward diabetes.

As already mentioned, iodomethane has reactivity towards distinct functional groups, which can be advantageous in that it can effect different modifications under appropriate conditions using the pH memory effect; and thus, the global electrostatic character of the protein molecule can be altered to different extents. The pH memory effect of lyophilized proteins allows for the selective chemical modification of ionizable functional groups in protein because it is possible to maintain in the lyophilized state the intrinsic nucleophilicity of the group, its microenvironment, and its ionization state, as there are no dynamic equilibria between ionization states.

Iodomethane modifies amino groups in protein to their quaternary state, which results in the preservation of positive charge for these groups over a wide pH range. Amino groups are involved in ionic interactions (i.e., salt bridges) with carboxylate groups, and these interactions play a role in maintaining protein structure. Mavri and Vogel (1994) have shown that the interaction energy of a salt linkage decreases with each level of methylation of the amino group. A maximum decrease of 25% upon quaternization of amino groups was calculated, which was not deleterious to the protein structure. The authors have suggested that such a subtle change in interaction energy in lysyl residues may modulate physiological function. For ribonuclease A, iodomethane modifies carboxylate salts to the neutral ester functionality at a degree of 37.4 ± 7.6% at LpH 7, which maintains at least 60% of the negative charges on the protein surface. When ribonuclease A is modified *in vacuo* at low LpH values with iodomethane the amino groups react minimally, as their pK_a values are high (6.8-8.0 and 10.4-11.1 for α- and ε-amino groups, respectively), and the carboxylate salt groups react to a small extent (between 14.2 ± 7.3 at LpH 3 and 25.3 ± 10.2 at LpH 6). Thus, chemical

modification of protein with iodomethane at low LpH will preserve at least 75-85% of the electrostratic global character of the protein molecule. Moreover, the chemical modification with iodomethane is small and preservation of the protein structure is likely. A change in steric bulk in the vicinity of the chemically modified group may perturb protein structure and consequently alter protein function. Methylation represents a small change in steric interactions and the effect of modification on the protein structure and function is presumably small. As mentioned above, Stewart *et al.* (1998) have studied the catalytic activity of α-chymotrypsin after *in vacuo* reaction with iodomethane. The enzyme was shown to retain activity if the active site is lyoprotected with the appropriate ligand. Protection or molecular imprinting and kinetic studies are necessary to further characterize the *in vacuo* esterification (with alcohol/HCl) and methylation (with iodomethane at low LpH values) reactions. (Histidine was found to react with iodomethane over a wide range of LpH values, and since it is required for activity in serine proteases, the active site of these enzymes must be protected by ligands during the *in vacuo* reaction. NMR spectroscopy can be used to establish the amount of ligand required to effect no reaction with histidine residues with iodomethane; i.e., maintainance of full enzymatic activity.) Acid proteases would be good test proteins for activity assays, since they are active at low pH values and the carboxyl methyl esters are stable under these conditions.

The preferential *in vacuo* esterification of carboxyl groups in lyophilized protein at low LpH values by iodomethane allows for the limitation of the extent of modification of these groups simply by choosing the appropriate reaction conditions (e.g., reaction time, LpH). Similarly, for the *in vacuo* esterification of carboxyl groups using methanol

and acid catalyst, the extent of modification can be controlled by choosing the reaction time and the amount of acid catalyst used (and even the LpH up to 6; higher LpH conditions may result in an increased degree of intra- and intermolecular crosslinking due to the presence of a higher proportion of unprotonated amino groups). The site-specific chemical modification of carboxyl groups in proteins is somewhat difficult to achieve, as is differentiation between aspartyl and glutamyl residues for the purpose of studying the roles of carboxyl groups in the structure and function of proteins (Lundblad, 1995a). However, it may be possible to utilize the nonaqueous methodology presented herein to effect site-specific or differential modification of carboxyl groups by fine-tuning the reaction conditions. For example, ^{13}C-methylation of ribonuclease A at LpH 3 almost exclusively resulted in the significant but small degree of derivatization of aspartyl residues (Figure 3-2).

The results of the present investigation of the *in vacuo* chemical modification of protein carboxyl groups show that it is possible to form stable activated derivatives of carboxyl groups in proteins with relatively small amounts of reagents compared to the high concentrations of activating reagents and nucleophiles which are required for aqueous modifications. The latter are generally required because activated carboxyl groups have short lifetimes in aqueous media, and thus, a high concentration of reagents is required to achieve extensive modification of these groups in proteins (Means and Feeney, 1971; Imoto and Yamada, 1989 and 1997; Lundblad, 1995a). Activated carboxyl groups in proteins can be coupled with nucleophiles to form derivatives that are stable in aqueous media. Coupling on a group that is negatively or positively charged or neutral can alter the physicochemical properties of the protein and may lead to interesting

structure-function relationship studies. Taralp (1997, Doctoral Thesis) esterified protein in the conventional manner by suspension in methanol in the presence of acid catalyst, then reacted with ethylenediamine, followed by *in vacuo* methylation to form a permanent positive charge on the carboxyl groups (via quaternization of the amine) of α-chymotrypsin (with derivatization at the C-terminus being of particular interest in a C-terminal identification strategy).

The SDS-PAGE result with methyl- or ethyl-esterified ribonuclease A showed evidence for covalent crosslinking between neighbouring protein molecules. A small proportion of higher molecular weight material compared to native protein was observed. The formation of intermolecular crosslinkages is likely to result from a nucleophilic unprotonated amino group (e.g., lysyl residues) displacing the methoxyl group of a carboxyl methyl ester. Although the number of unprotonated amino groups is small at low LpH, evidence for their presence in the lyophilized protein structure has been observed in the present study, and was provided from the *in vacuo* methylation of human albumin with iodomethane (see Figure 3-3 human albumin, LpH 3, 4 or 5).

An alternative explanation for the observation of ^{13}C-trimethylamino resonances of lysyl residues in human albumin at LpH 3 is the following. In aqueous medium, there is a dynamic ionization equilibrium between unprotonated and protonated ionization states of an amino group, which shifts toward the unprotonated state prior to nucleophilic attack on iodomethane. In the gas phase, a dynamic equilibrium between ionization states of amino groups does not exist, or is very slow; hence, a protonated amino group in a lyophilized protein at low LpH is the only form of amino group that is available to react with iodomethane, albeit at a slower rate than an unprotonated amino group, by

transferring a proton from the protonated amino group to an anionic counterion. The products of this reaction are the monomethyl amino derivative of the amino acid residue (which reacts further with iodomethane to form the trimethyl amino derivative via the dimethyl derivative), hydroiodic acid (HI) and HX (where X is a counterion such as Cl^-, HPO_4^- from phosphate buffer in the protein preparation, or a carboxylate anion within the protein structure itself or within a neighboring protein molecule). Hence, a likely scenario is that the protonated amino group would transfer its proton to a counterion *in vacuo*, rendering a nucleophilic amino group that can attack an activated carboxyl methyl ester (formed in the reaction of protein with methanol and acid catalyst) to form an amide crosslinkage and methanol. If proton transfer occurs under the *in vacuo* conditions, it is not unreasonable to suggest that a crosslink may form at an intermolecular salt bridge with the concomitant loss of water. This avenue of research is currently being pursued by Ms. Brigitte Simons.

The extent of crosslinking is likely to be protein dependent and would be an interesting study for future work. Moreover, the formation of activated carboxyl groups to a high degree, as observed with the *in vacuo* esterification reaction is an invitation to the study of the optimization of intermolecular crosslinking of desired proteins for the purpose of increasing their stability (Lundblad, 1995 (b); Dugas, 1996; Kluger in Creighton, 1997) in the hope of employing them in biotechnological industrial applications. Optimal mixing of native and activated protein at a moderate pH or in nonaqueous hydrophobic media may be two experimental routes to this end. A higher degree of crosslinking (relative to that observed for ^{13}C-methyl- or ^{13}C-ethyl-esterified ribonuclease A at LpH 6) is believed to be attainable because ^{13}C-methyl-esterified

ribonuclease A at LpH 10 was observed to be virtually impossible to solubilize, even in 10 M urea, which did not allow for solution ^{13}C NMR spectroscopic analysis or molecular weight determination by SDS-PAGE. A higher degree of crosslinking in the latter sample was hypothesized and presumably resulted from the greater proportion of unprotonated amino groups available to nucleophilically attack the electropositive carbonyl carbon atom of carboxyl methyl esters with the subsequent displacement of methoxyl groups. Osbahr (1982) has shown that the esterification of fibrinogen carboxyl groups with the group-specific trimethyloxonium ion resulted in the polymerization of the modified protein molecule, with a minimum number of methyl groups being incorporated into the fibrinogen molecule; the polymeric material resembled the physiologically formed fibrin clot. Intermolecular covalent crosslinking between specific amino acid residues in protein also has been used in the study of protein-protein interactions (Brinkley, 1992). If intermolecular crosslinking of protein is undesirable, lyoprotectants could be used to limit the interaction of molecules from interacting as much with each other (Dabulis and Klibanov, 1993). The latter may even enhance the extent of chemical modification of protein carboxyl groups and is another desired avenue for future research.

Furthermore, the extent of methylation seems to depend on how the protein is packed in the lyophilized state. A protein cake may not expose all surface functional groups, and thus lead to a lower extent of reaction. Additional studies are required to observe the maximum extent of methylation on lyophilized protein powder over a large surface area at each LpH value. Another variable that requires further study is the lyophilization volume. It may be possible to increase the extent of methylation of a

protein if it is lyophilized from a 0.1 mg/mL solution, which would presumably keep the molecules separated and result in a higher percentage of surface-exposed reactive groups. A protein lyophilized from such a dilute solution should also promote intramolecular crosslinking (Lundblad, 1995b). Taralp (Doctoral Thesis, 1997) found that the only factor which could be controlled, and which drastically affected the extent of crosslinking, was the lyophilization volume. Taralp (Doctoral Thesis, 1997c) also observed that crosslinking was significant when couplings with nucleophiles (e.g., a fluorescent amine in *N,N*-dimethylformamide) were carried out using carboxyl group activating reagents such as carbodiimides or anhydrides. Thus, if undesired, crosslinking can pose a problem to the protein chemist.

In summary, *in vacuo* esterification reactions were investigated in the present study. Preferential esterification of protein carboxyl groups was observed by reacting the lyophilized protein at low LpH values with ^{13}C-iodomethane. The *in vacuo* reaction of lyophilized protein with ^{13}C-methanol/HCl or ^{13}C-ethanol/HCl resulted in the selective esterification of protein carboxyl groups. Hence, *in vacuo* esterification with alcohol/HCl is the obvious procedure to use for specific methylation of lyophilized protein carboxyl groups. Although 100% methyl-esterification of carboxyl groups is difficult to achieve, the extent attainable (60-70%) is as good as the conventional method for acid-catalyzed esterification by suspension of the protein in acidic alcohol (Ram and Maurer, 1959; Broomfield *et al.*, 1965; Means and Feeney, 1971). Hence if extensive esterification of carboxyl groups is desired, then *in vacuo* reaction with methanol/HCl is the method of choice. Ethanol/HCl esterified carboxyl groups in the lyophilized protein to a much lesser extent (25%) than did methanol/HCl. It appears that not all carboxyl groups in the

lyophilized protein are freely accessible and the bulkiness of the alcohol is an important factor in determining the extent of esterification. *In vacuo* esterification with ethanol/HCl could be very useful in cases where it is desirable to form a more stable ester derivative, or to introduce a more stable ester derivative containing a ^{14}C, ^{13}C, CD_2, CD_3, CT_2 or CT_3 label into the protein. Reaction with iodomethane was useful in that the pH memory effect of lyophilized proteins (Zaks and Klibanov, 1988; Taralp, Doctoral Thesis, 1997) was confirmed; but the results show that when alkylating amino groups (Chapter 2; Vakos *et al.*, 2000), *in vacuo*-synthesized esters of carboxyl groups will be formed and should be removed. Moreover, it is much easier to control the extent of reaction with iodomethane, so in cases where a low extent of esterification of carboxyl groups is desired, reaction with iodomethane would be preferable.

It was possible to distinguish the three types of protein carboxyl group derivatives, namely the α-, γ-, and δ-carboxyl ^{13}C-methyl and ^{13}C-ethyl esters from their distinct chemical shifts by ^{13}C-NMR spectroscopy. Within the class of γ- or δ-carboxyl esters, the variation in the chemical shifts of individual esterified side-chains of Asp or Glu residues was less than 0.05 ppm, so that resonances from individual residues could not always be resolved at 9.4 Tesla. In contrast, the chemical shift(s) of the C-terminal α-carboxyl group(s) showed a strong dependence on the nature of the C-terminal amino acid and sequence (i.e., penultimate amino acid residue). It was possible to differentiate two forms of insulin that differed only in the penultimate amino acid of the B chain. Incorporation of a ^{13}C-label into α-carboxyl groups provides a new approach for determining if a pure protein consists of more than one polypeptide chain, or if a purified protein contains a significant amount of a contaminating protein of similar size, which

would be difficult to detect by SDS-PAGE. It should be possible to make a tentative identification of the C-terminus of a protein by comparing the chemical shift of the α-carboxyl methyl ester observed in the protein with that of the C-termini in standard peptides.

To our knowledge, iodomethane or methanol/HCl (or ethanol/HCl) have not been employed to effect *in vacuo* esterification of lyophilized protein carboxyl groups; however, iodomethane has been used to modify the pentapeptide Leu-enkephalin in the vapour phase (Vath *et al.*, 1988). The extent of modification by the *in vacuo* esterification reaction using iodomethane or alcohol/HCl can be controlled by careful choice of the reaction conditions, such as reaction time, LpH (reaction with iodomethane), amount of acid catalyst (reaction with alcohol/HCl), etc. For the *in vacuo* esterification reaction with ^{13}C-iodomethane, the extent of derivatization was directly related to the protein solution structure prior to lyophilization, which reflects the solution pH from which the protein was lyophilized. The physicochemical factors governing the reactivity of protein functional groups in the *in vacuo* nonaqueous environment reflect the pK_a and surface accessibility of reactable functional groups in the native protein. The *in vacuo* approach to chemical modification of protein allows for the formation of derivatives with different degrees of modification, which can potentially be used to study novel structure-function relationships. Moreover, it affords the opportunity to use small amounts of reagents, which allows for the introduction of isotopic labels (^{13}C, ^{14}C, etc.) without disruption of protein structure (Taralp, Doctoral Thesis, 1997; Stewart *et al.*, 1997). The present work was carried out with a 9.4 Tesla instrument. It is expected that

modern higher field spectrometers would provide significant increases in both sensitivity and spectral resolution.

5. Acknowledgement

This research was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

6. References

Acharya A. S. and Vithayathil, P. J. (1975) *Int. J. Pept. Protein Res.*, **7**: 207-219.

Beckwith, A. L. J. (1970) In *The Chemistry of Amides* (Zabicky, J., Ed.), p. 106.

Bendall, M. R., Doddrell, D. M., Pegg, D. T., and Hull, W. E. (1982). In *DEPT, Bruker*

Blackburn, S., Hemphill Carter, E., and Phillips, H. (1941) *Biochem. J.*, **35**: 627-639.

Bollag, D. M. and Edelstein, S. J. (1991) Protein Methods, Wiley-Liss, New York, p.100.

Breitmaier, E., and Voelter, W. (1987). In *Carbon-13 NMR Spectroscopy: High-Resolution Methods and Applications in Organic Chemistry and Biochemistry*, 3rd ed., VCH Publishers, New York. (a) p. 109; (b) p. 208; (c) p. 228.

Brinkley, M. (1992) *Bioconjugate Chem.*, **3**: 2.

Broomfield, C. A., Riehm, J. P., and Scheraga, H. A. (1965) *Biochemistry*, **4**: 751-759.

Chibnall, A. C., Mangan, J. L., and Rees, M. W. (1958) *Biochem. J.*, **68**: 114.

Clairmont, F. (1998) In *Structure of the Insecticidal Crystal Protein from Bacillus thuringiensis var. kurstaki HD-73.* Ph.D. Thesis, Department of Chemistry, University of Ottawa, Ottawa, ON.

Creighton, T. E. (1993). *Proteins: Structure and Molecular Properties*, 2nd ed., W.H. Freeman and Company. (a) p. 5; (b) p. 6

Dabulis, K. and Klibanov, A. M. (1993) *Biotech. Bioeng.*, **41**: 566-571.

Darbre, A. (1986). *Practical Protein Chemistry* (Darbre, A., ed.), Wiley, New York. (a) p.81.

Dayhoff, M. O., ed. (1972) Atlas of Protein Sequence and Structure, Vol. 5, The National Biomedical Research Foundation, Silver Springs, Maryland.

Dugas, H. (1996) *Bioorganic Chemistry: A Chemical Approach to Enzyme Action*, 3rd edn., Springer, pp. 230-233.

Duggleby, R. G. and Kaplan, H. (1975) *Anal. Biochem.*, **65**: 346-354.

Ferry, J. D. (1948) *Adv. Protein Chem.*, **4**: 1.

Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd ed., W.H. Freeman and Company, New York.

Fraenkel-Conrat, H. and Olcott, H. S. (1945) *J. Biol. Chem.*, **161**: 259-268.

Glasoe, P. K. and Long, F. A. (1960) *J. Phys. Chem.*, **64**: 188-190.

Harris, D.C. (1987) *Quantitative Chemical Analysis*, 2nd ed. (a) pp., 729, 730, 732; (b) p. 731.

Harris, R. K. (1983) *Nuclear Magnetic Resonance Spectroscopy: A Physicochemical View*, Pitman Publishing Inc., pp.67-68, pp.85-93.

Halpin, M.I. and Richardson, T. (1985) *J. Dairy Sci.*, **68:** 3189-3198.

Herriott, R. M. (1947) *Adv. Protein Chem.*, **3:** 169.

Hoare, D. G. and Koshland, D. E. (1966) *J. Biol. Chem.*, **242:** 2447.

Imoto, T. and Yamada, H. (1989) In *Protein Function* (Creighton, T. E., Ed.), IRL Press, Oxford.

Imoto, T. and Yamada, H. (1997) In *Protein Function* (Creighton, T. E., Ed.), IRL Press, Oxford, pp. 279-316.

Jencks, W. P. (1958) *J. Am. Chem. Soc.*, **80:** 4585-4588.

Johnson, B. A. and Aswad, D. W. (1985) *Biochemistry*, **24:** 2581-2586.

Kaplan, H., Stevenson, K. J., and Hartley, B. S. (1971) *Biochem. J.*, **124:** 289-299.

Kim, S. and Paik, W. K. (1976) *Experimentia*, **32:** 982-984.

Kluger, R. (1997) In *Protein Function* (Creighton, T. E., Ed.), IRL Press, Oxford, pp. 185-213.

Kruizinga, Strijtveen, and Kellogg (1981) *J. Org. Chem.*, **46:** 4321.

Link, T. P. and Stark, G. R. (1968) *J. Biol. Chem.*, **243:** 1082-1085.

Lundblad, R. L. (1995) *Techniques of Protein Modification*, CRC Press, Boca Raton. (a) Ch. 14; (b) Ch. 15.

Magdassi, S., ed. (1996) *Surface Activity of Proteins, Chemical and Physiochemical Modifications*, Marcel Dekker, Inc., p.28.

March, J. (1985) In *Advanced Organic Chemistry*, Third Edition, Wiley-Interscience, New York. (a) p.335 and 349; (b) p.340; (c) p.351; (d) p.353

Mavri, J. and Vogel, H. J. (1994) *Proteins: Structure, Function and Genetics*, **18:** 381-389.

McCumber, L. J. and Clem, L. W. (1976) *Biochim. Biophys. Acta*, **446:** 536-541.

Means, G.E. and Feeney, R.E. (1971) *Chemical Modification of Proteins*, Holden-Day, pp.139-148.

Osbahr, A. J. (1982) *Thromb. Haemost.*, **48:** 226-231.

Paik, W. K. and Kim, S. (1980) *Protein Methylation*, Wiley-Interscience, New York. (a) p.47; (b) p.202; (c) p.214; (d) p.215

Parsons, S. M., Jao, L., Dahlquist, F. W., Borders, C. L., Jr., Groff, T., Racs, J., and Raftery, M. A. (1969) *Biochemistry*, **8:** 700.

Parsons, S. M. and Raftery, M. A. (1969) *Biochemistry*, **8:** 4199.

Paterson, A. K. and Knowles, J. R. (1972) *Eur. J. Biochem.*, **31:** 510.

Perutz, M. F. (1978) *Science*, **201:** 1187.

Ram, J. S. and Maurer, P. H. (1959) *Arch. Biochem. Biophys.*, **85:** 512.

Reisse, J. (1982) In *The Multinuclear Approach to NMR Spectroscopy* (Lambert, J. B. and Riddell, F. G., Eds.), D. Reidel Publishing Company, Ch. 4.

Riehm, J. P., Broomfield, C. A., and Scheraga, H. A. (1965) *Biochemistry*, **4**: 760-771.

Robinson, A. B. (1974) In *Current Topics in Cellular Regulation* (B. L. Horecker and E. R. Stadman Eds.), Vol. 8, Academic Press, New York, pp. 247-295.

Rupley, J. A., Gratton, E., and Careri, G. (1983) *TIBS*, 18-22.

Rutherford, H. A., Patterson, W. I., and Harris, M. (1940) *J. Res. Nat. Bur. Stds.*, **25**: 451-458.

Steele, J. C. H., Jr., and Nielson, T. B. (1978) *Anal. Biochem.*, **84**: 218.

Stewart, N.A.S., Taralp, A., and Kaplan, H. (1997) *Biochem. Biophys. Res. Commun.*, **240**: 27-31.

Strickley, R. G. and Anderson, B. D. (1996) *Pharm. Res.*, **13**: 1142-1153.

Taralp, A. (1997) In *Aqueous and Nonaqueous Chemical Approaches to Elucidating Structure and Function in Proteins*. Ph.D. Thesis, Department of Chemistry, University of Ottawa, Ottawa, ON. (a) pp.56-58; (b) p.106; (c) p.119.

Taralp, A. and Kaplan, H. (1997) *J. Protein Chem.* **16**: 183-193.

Vakos, H. T., Burton, G. W., and Kaplan, H. (1997) *Biochim. Biophys. Acta*, **1342**: 103-108.

Vakos, H. T., Kaplan, H., Black, B., Dawson, B., and Hefford, M. A. (2000) *J. Protein Chem.*, **19**: 231-7.

Vath, J. E., Zollinger, M., and Biemann, K. (1988) *Fresenius Z. Anal. Chem.*, **331**: 248-252.

Vollhardt, K. P. C. (1987) *Organic Chemistry*, W. H. Freeman and Company, New York. (a) p. 957-958; (b) pp. 806-807

Wilcox, P. E. (1967) In *Meth. Enzymol.*, **11**: 605.

Wissmann, A. and Hillen, W. (1991) In *Meth. Enzymol.* (Sauer, R. T., Ed.), Academic Press, Ch. 18, p. 208.

Woodward, R. B., Olofson, R. A., and Mayer, H., (1961) *J. Am. Chem. Soc.*, **83**: 1010.

Woodward, R. B., Olofson, R. A. (1966) *Tetrahedron*, Suppl. **7**: 415.

Young, N. M. and Kaplan, H. (1989) In *Protein Function* (Creighton, T. E., Ed.), IRL Press, Oxford.

Zaks, A. and Klibanov, A. M. (1988). *J. Biol. Chem.*, **263**: 3194-3201.

Appendix 3-1

Table 3-5A. Raw Difference Titration Data for the Estimation of the Extent of Modification of Carboxyl Groups in Ribonuclease A

Trial	Reaction conditions RNase A, no buffer	μL 0.05 N NaOH Forward titration†	μL 0.05 N HCl Backward titration†	Extent of methylation* % Forward titration	Extent of methylation* % Backward titration
	LpH 3, MeI				
1-1		215.5	195.5		13.3
1-2		237.5	236.5		-
2-1		216	184.5		22.5
2-2		235.5	205.5		4.9
3-1		200	192		16.2
average P-BL		*201.5*	*211.4*		
average W-BL		*77*	*92*		
	LpH 4, MeI				
1-1		213.5	181.5		25.0
1-2		209.3	189		18.8
2-1		219.5	206		4.5
2-2		204.0	207.5		3.3
3-1		264.5	161		12.9
3-2		205	196		31.7
4-1		181	173.5		
average P-BL		*201.5*	*211.4*		
average W-BL		*77*	*92*		
	LpH 5, MeI				
1-1		281.5	253		-
1-2		173	182		24.6
2-1		194	182.5		24.2
2-2		190.5	177		28.8
3-1		191	170		34.7
3-2		195	187.5		20.0
4-1		184	171		33.8
average P-BL		*201.5*	*211.4*		
average W-BL		*77*	*92*		

	LpH 6, MeI			
1-1		191	187	20.4
1-2		177.5	174.5	30.9
2-1		188	173	32.2
2-2		208	183	23.8
3-1		211.5	195	13.7
3-2		194.5	194	14.6
4-1		185	161.5	41.8
average P-BL		*201.5*	*211.4*	
average W-BL		*77*	*92*	
	LpH 7, MeI			
1-1		192	172.5	32.6
1-2		169.5	157	45.6
2-1		176	172.5	32.6
2-2		190	173	32.2
3-1		167	166	38.0
3-2		173.5	174.5	30.9
4-1		163.5	151.5	50.2
average P-BL		*201.5*	*211.4*	
average W-BL		*77*	*92*	
	LpH 10, MeI			
1-1		151	131	67.3
1-2		145.5	139	60.6
2-1		138	130	68.2
2-2		162	159	43.9
3-1		153	153	48.9
average P-BL		*201.5*	*211.4*	
average W-BL		*77*	*92*	

	LpH6, MeOH/HCl 25:5				
1-1		203	287	-	-
1-2		178	266	18.9	4.4
2-1		187	260	11.6	7.7
2-2		179	254	18.1	11.0
average P-BL		*201.5*	*274*		
average W-BL		*77*	*92*		
	LpH6, MeOH/HCl 25:10				
1-1		195	249.5	15.5	13.5
1-2		192	240	17.7	18.7
2-1		178	215	18.9	20.5
2-2		175	234	21.3	-
average P-BL		*216.7, 201.5*	*274*		
average W-BL		*77*	*92*		
	LpH6, MeOH/HCl 25:20				
1-1		162.5	194.0	31.3	34.1
1-2		153	181	39.0	42.5
2-1		142	165	47.8	52.8
2-2		146	216	44.6	19.9
average P-BL		*201.5*	*246.8*		
average W-BL		*77*	*92*		

†forward reaction titrated between pH 3.00-5.50; backward reaction titrated between pH 5.50-3.00

* $\frac{\text{P-BL} - \text{S}}{\text{P-BL}} \times 100\%$

*normalized to P-BL {i.e., [P-BL – S (x mg P-BL avg./mg S avg.)]/[P-BL – BL]}

***Based on:**

$$\frac{[\#\mu\text{L P-BLANK (+ W-BLANK)}] - [\#\mu\text{L SAMPLE (+ W-BLANK)}]}{[\#\ \mu\text{L P-BLANK (+ W-BLANK)}] - \text{W-BLANK}} \times 100\%$$

$$\frac{\text{P-BL} - \text{S}}{\text{P-BL}} \times 100\%$$

*See Table 3-5; P-BL=native protein blank, (+W-BL)=water blank inclusive in the P-BL or S. S=sample. W-BL=water blank as determined in the absence of P-BL or S.

Appendix 3-2

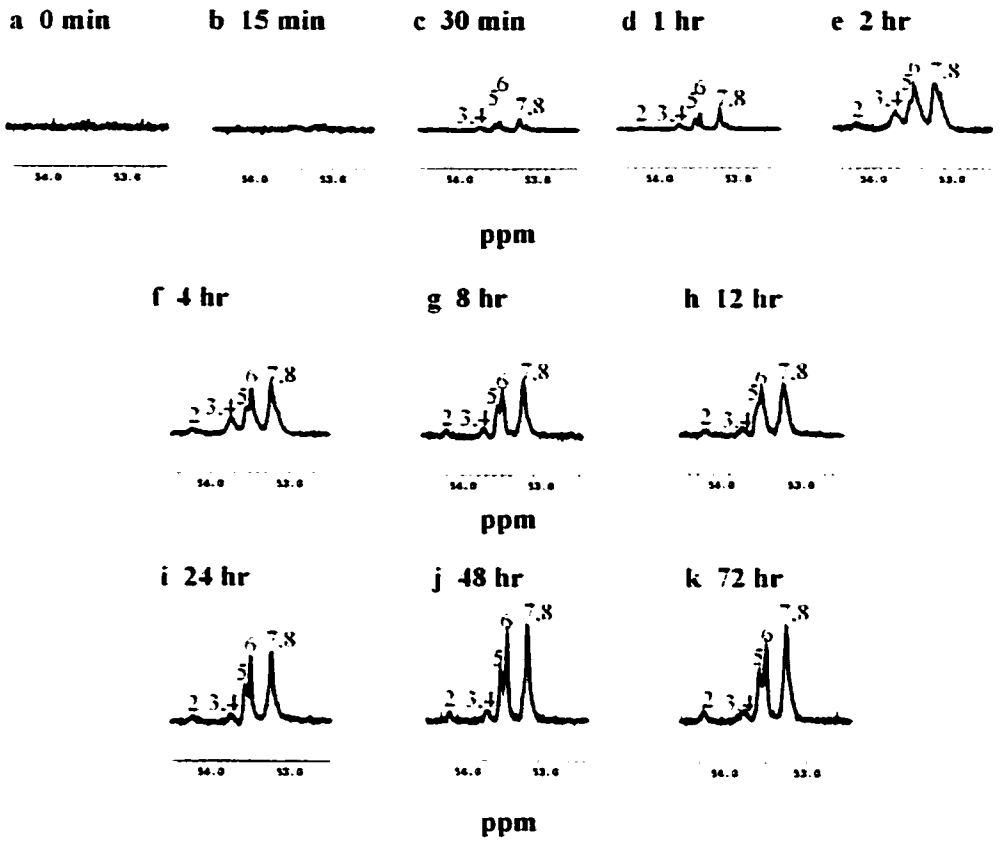


Figure 3-1A. Expanded ^{13}C NMR spectra of ^{13}C-methylated ribonuclease A (LpH 5; reacted *in vacuo* with ^{13}C-iodomethane at 75°C) at various times.

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1) unidentified small molecule (56.21 ppm); (2) unidentified α-$^{+}NMe_3$-impurity (54.23 ppm); (3) $Me_3N^{\varepsilon+}$-Lys (53.76 ppm); (4) $C_\alpha OOMe$ (53.71 ppm); (5) $C_\gamma OOMe$ (53.56 ppm); (6) $C_\gamma OOMe$ (53.48 ppm); (7) $Me_3N^{\alpha+}$-Lys (53.24 ppm); (8) $C_\delta OOMe$ (53.23 ppm).

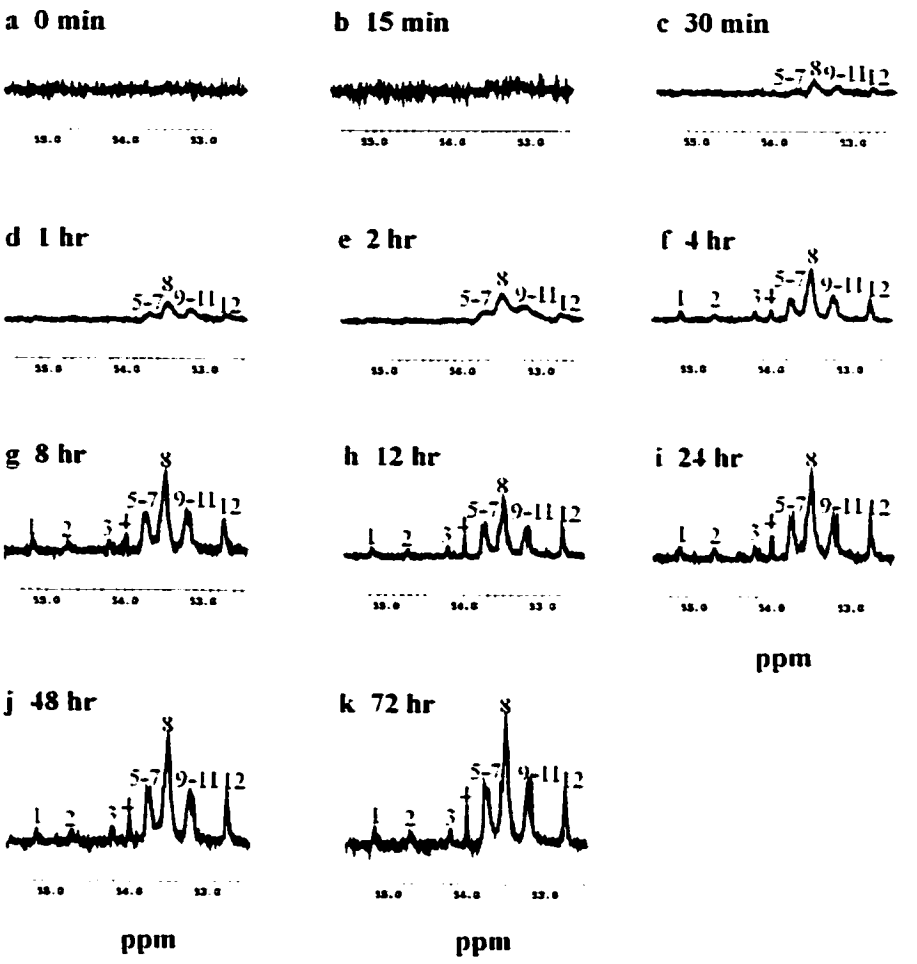


Figure 3-2A. Expanded ^{13}C NMR spectra of ^{13}C-methylated α-chymotrypsin (LpH 5; reacted *in vacuo* with ^{13}C-iodomethane at 75°C) at various times.

Peak resonances correspond to the ^{13}C-enriched methyl groups in:

(1), (2), (3) unidentified α-$^{\frown}$N**Me**$_3$, impurities (55.16 ppm, 54.72 ppm, and 54.19 ppm); (4) C_αO**OMe** (53.99 ppm; C-terminal Asn-245); (5) C_αO**OMe** (53.77 ppm); (6) **Me**$_3$N$^{\varepsilon+}$-Lys (53.74 ppm); (7) C_αO**OMe** (53.72 ppm); (5, 7) C-termini: Leu-13, Tyr-146; (8) C_γO**OMe** (53.50 ppm); (9) half-cystine-$^{\frown}$N$^{\alpha}$**Me**$_3$; (10) C_δO**OMe** (53.21 ppm); (11) **Me**$_3$N$^{\alpha+}$-Ile (53.17 ppm); (12) **Me**$_3$N$^{\alpha+}$-Ala (52.74 ppm).

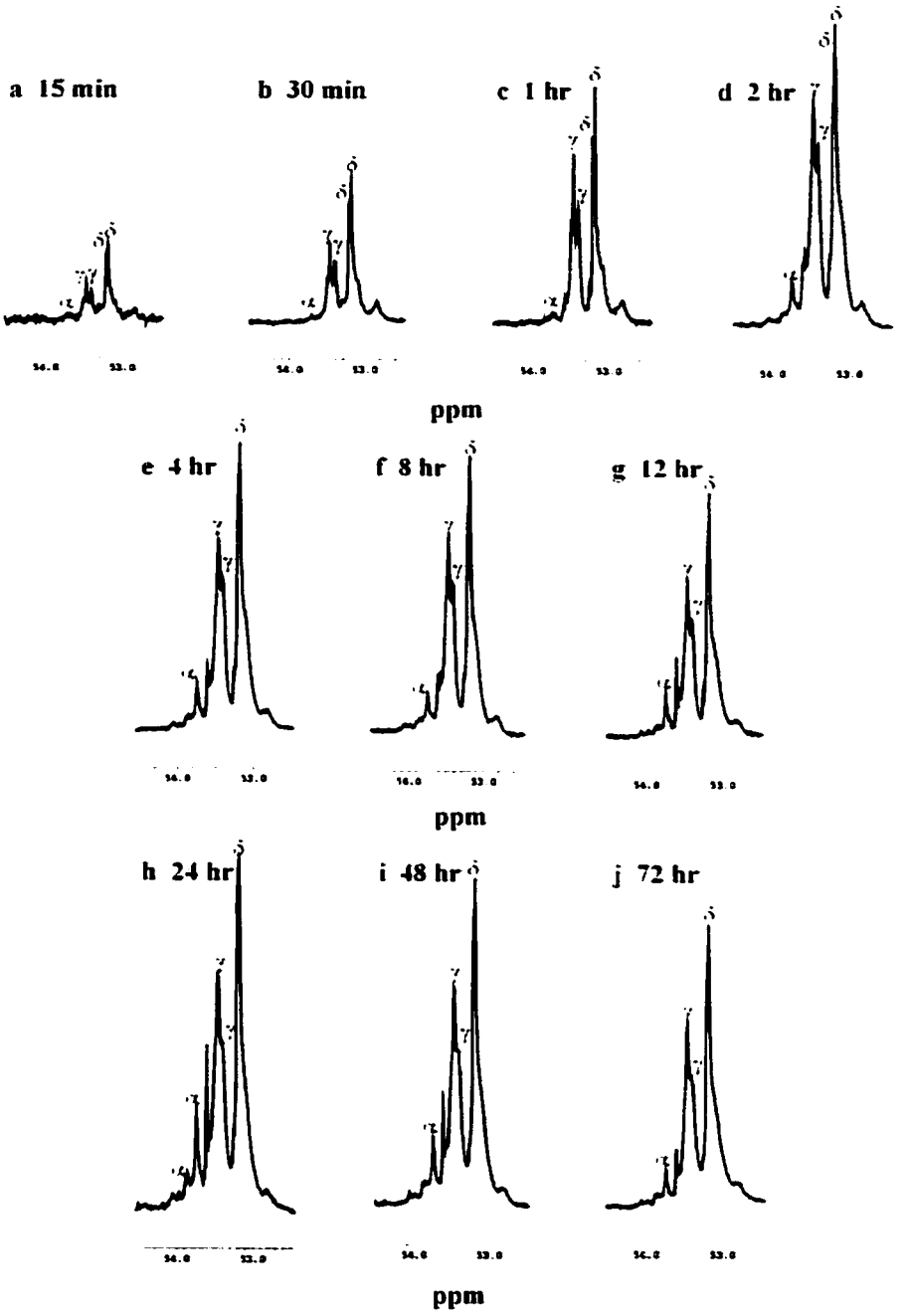


Figure 3-3A. Expanded ^{13}C NMR spectra of *in vacuo* ^{13}C-methyl- esterified (25:10 $^{13}MeOH/HCl$) ribonuclease A (LpH 3, 75°C) at various times.

The ^{13}C-resonances corresponding to the α-, γ-, and δ-carboxyl methyl esters are: 53.74 (α), 53.61 (methylated small molecule), 53.46 (γ), 53.41 (γ), 53.21 (δ), and 53.18 ppm (δ).

Chapter 4

The Use of ^{14}N NMR Spectroscopy To Study Trimethylated Amino Groups in Methylated Protein Hydrolysates

1. Introduction

Novel applications of existing NMR techniques in the area of protein characterization are of interest to the protein chemist. Taralp (1997; Doctoral Thesis) has demonstrated the use of ^{14}N NMR spectroscopy to identify or observe N^{α}- and N^{ε}-amino groups in proteins and amino acids. The present study is an extension and verification of the preliminary investigation of Taralp. ^{14}N NMR spectroscopy is used to characterize the trimethylamino derivatives of several *in vacuo* methylated model proteins; methylated amino acid standards were prepared *in vacuo* using the technique of Taralp and Kaplan (1997) and by the published protocol for exhaustive methylation (Greenstein and Winitz, 1961). Although ^{14}N nuclei have a high natural isotopic abundance of 99.6% (Witanowski and Webb, 1972a), other nuclear (NMR) properties have limited the use of ^{14}N NMR spectroscopy on protein systems as a structure-function probe. These are low to average receptivity, low resonance frequencies, and broad resonance lines that arise from very short relaxation times, and are a result of the large quadrupole moment of the ^{14}N nucleus (Witanowski and Webb, 1972a).

^{14}N-Linewidths have been studied as a function of pH for several amino acids, related molecules, and protein amino groups (Tzalmona and Loewenthal, 1974; Cohen *et al.*, 1975; Richards and Thomas, 1975; Zhang *et al.*, 1994; Jenks, 1971). Protonation states and proton exchange processes in aqueous solution have been shown to affect substantially the ^{14}N lineshapes and chemical shifts (Cooper *et al.*, 1973; Tzalmona and

Loewenthal, 1974; Cohen *et al.*, 1975; Blomberg *et al.*, 1976; Sheinblatt and Gutowsky, 1976).

Tzalmona and Loewenthal (1974) have reported the ^{14}N NMR for glycine as a function of pH, and Richards and Thomas (1975) have reported studies of histidine and histamine. Cohen *et al.* (1975) have studied the ^{14}N NMR linewidth versus pH dependences of several amino acids and related molecules. They have shown that the linewidth pH profile of each amino acid is unique. Interestingly, linewidth minima were shown to occur between pH 6 and 8; and substantial changes of linewidth were shown to occur with changes in pH. In the case of glycine, the transverse relaxation times and linewidths at low pH were affected by proton exchange, while at pH values greater than 6, the relaxation times and linewidths were a weighted average corresponding to the proportion of zwitterionic and anionic species at a given pH value (Tzalmona and Loewenthal, 1974). Linebroadening was greatest at either pH extreme.

Nuclei with spins greater than ½ have their NMR spectra broadened by interaction of the nuclear quadrupole moment with a fluctuating electric field gradient at the nucleus due to molecular motion. The results of Cohen *et al.* (1975) are as expected, since the change of structure of the amino acids with pH would modulate the electric field gradients that the nitrogen nuclei of the amino groups experience. These changes, along with changes in exchange rate between species, can affect the nitrogen nuclear quadrupole relaxation time, and thus, the ^{14}N NMR linewidth. Indeed, the relaxation time of a nucleus with spin greater than ½ can be a very sensitive function of the molecular environment.

Zhang and coworkers (1994) have shown that severe peak broadening occurs as the pH of the sample (e.g., ^{14}N NMR of L-lysine, N^{ε}-monomethyl-L-lysine, and N^{ε}-dimethyl-L-lysine) exceeds the pK_a value of the amino group. In contrast, the tetraalkylammonium nitrogen of N^{ε}-trimethylamino-L-lysine in aqueous solution showed one narrow resonance (3 Hz) due to the tetrahedral electronic distribution about the quaternary nitrogen. This high symmetry around the ^{14}N nucleus is indicative of a longer T_2^* relaxation time for the ^{14}N nucleus since $\Delta\nu_{1/2}$ is proportional to $1/T_2^*$, and hence, is responsible for the sharper linewidths compared to more unsymmetrical ^{14}N nuclei of amino groups. Moreover, the chemical shift of the quaternary trimethylated nitrogen did not change in the course of the pH titration, which indicated that there was no titratable proton. This is consistent in that quaternary alkylated nitrogens should remain positively charged at all pH values; as expected then, the ^{14}N resonance retained the same linewidth and chemical shift over a wide pH range (Zhang *et al.*, 1994).

^{14}N NMR spectroscopy has been applied to the field of peptide hormones; chemical shift positions and linewidths of quaternary nitrogens of the amino-terminal group resonance of the brain neurotransmitter pentapeptides, Leu- and Met-enkephalins, and enkephalin-related peptide fragments in aqueous solution have been studied (Gerothanassis *et al.*, 1987; Karayannis *et al.*, 1990). Karayannis *et al.* (1990) have employed ^{14}N and ^{17}O NMR spectroscopy as structural probes of amino acids and have applied these tools to the study of the conformational properties of enkephalins.

NMR spectroscopy has the potential of being used as a probe for protein structure since relaxation processes of NMR-active nuclei are sensitive to molecular motions (Berglund *et al.*, 1995). Proteins tumble slowly in solution, but the side-chains of amino

acid residues are characterized by fast local motions (Breitmaier and Voelter, 1987a). The study of internal dynamics of proteins and other large molecules has been made possible through technical advances; relaxation measurements can be carried out that separate the two dynamic components (Berglund *et al.*, 1995). Unlike the usual employment of ^{13}C, ^{15}N, and ^{2}H nuclei for motional studies (Fersht, 1985), the use of the quadrupolar ^{14}N nuclei has been limited due to the problem of broad resonance lines.

Galanter and Labotka (1991) have studied the binding of [^{14}N]nitrate, a transition state analogue of the bicarbonate substrate, to the human erythrocyte anion transport protein, AE1, by measuring the increase of linewidth upon association. The interaction of the ^{14}N quadrupolar nucleus with a protein can broaden the solution resonance by affecting the transverse relaxation rate (Drakenberg and Forsen, 1983). Normally, the $^{14}NO_3^-$ resonance of free nitrate in solution is sharp because the ^{14}N nucleus has a symmetrical trigonal planar electronic environment. However, when it is bound to the erythrocyte protein, its linewidth at half-height becomes much larger due to the electronic asymmetry imposed on the ^{14}N nucleus by the binding pocket of the protein. The apparent linewidth obtained is thus a weighted average of free and bound nitrate, and as such, increases with increasing concentration of the protein. Further evidence that nitrate binds the erythrocyte protein was obtained by adding an AE1 specific competitive inhibitor, which reduced the linewidth to that of free nitrate, presumably by displacement of all bound nitrate. Similarly, bicarbonate and chloride were found to be competitive inhibitors of the AE1 specific $^{14}NO_3^-$ resonance line-broadening effect. Based on the concentration dependence of inhibition and by using a model of competitive inhibition, the K_D of bicarbonate binding to AE1 was estimated. Hence, the ^{14}N NMR nitrate

binding assay was shown to be a powerful tool for studying the structure of the AE1 binding site for the physiological substrate, bicarbonate.

Zhang *et al.* (1994) have carried out dynamic ^{14}N NMR experiments to probe the microenvironment of an N^{ϵ}-trimethylamino lysyl residue in cytochrome c and calmodulin. The amino groups of the proteins generally gave variable broad lines that were exaggerated with an increase in pH. Deprotonation of amino groups upon increasing the pH has been shown to account for the extensive linewidth (Gerothanassis *et al.*, 1987; Tzalmona and Loewenthal, 1974; Cohen *et al.*, 1975). The proteins studied were special in that both had one N^{ϵ}-trimethylamino lysyl residue present as a post-biosynthetic (post-translational) modification. This quaternary trimethylated ammonium group of lysine had a small and constant linewidth (23 Hz), and unlike the other amino groups of the protein, did not broaden with an increase in pH (i.e., did not show any pH dependency). Analysis with the free amino acid analogues, namely lysine, N^{ϵ}-methyl lysine, N^{ϵ}-dimethyl lysine, and N^{ϵ}-trimethyl lysine also followed the same trends; however, the linewidth of the free trimethyl lysine was an impressive 3 Hz. Hence, the trimethylamino lysyl residues of cytochrome c and calmodulin were observed as relatively narrow resonances. The latter made possible motional NMR studies of the region surrounding the lysyl residue. While Zhang *et al.* (1994) have demonstrated the utility of trimethylammonium groups of protein and ^{14}N NMR spectroscopy for the study of protein dynamics, the disadvantage of the work was that it required the proteins to be post-biosynthetically trimethylated.

In Chapter 2, the *in vacuo* methylation procedure of Taralp and Kaplan (1997; Taralp, 1997, Doctoral Thesis) was utilized to analyze N^{α}- and N^{ϵ}-amino groups in

protein by ^{13}C-NMR spectroscopy. The reaction of protein amino groups with iodomethane proceeds to the quaternary ammonium salt, and occurs, in general, with retention of protein structure and, in some cases, biological activity (Taralp, 1997, Doctoral Thesis). *In vacuo* methylated chymotrypsin was fully active when the protein was lyophilized in the presence of active site protecting ligands prior to reaction with iodomethane (Stewart, Taralp and Kaplan, 1997). Enzymes in nonaqueous environments have been shown to retain their structure and activity at temperatures as high as 100°C (Zaks and Klibanov, 1984; Zaks and Klibanov, 1988). The chemical modification procedure was advantageous over the work of Zhang *et al.* (1994) since the ability to trimethylate any amino group in any protein system negates the requirement for a biologically methylated protein system, and for this reason, could be used as a general structure-function tool. It follows, therefore, that advantage can be taken of ^{14}N NMR spectroscopy to: 1) monitor the reaction of amino functional groups in proteins with iodomethane; 2) identify and characterize the derivatized N^{α}- and N^{ϵ}-amino acid residues; and 3) obtain information on the local dynamics and microenvironment of the amino group from which the trimethyl ammonium derivative was originally prepared. While ^{13}C-NMR spectroscopy has been similarly applied to study N^{α}- and N^{ϵ}-amino groups in methylated proteins and amino acids (Taralp and Kaplan, 1997) employing costlier ^{13}C-iodomethane, the much larger chemical shift range (but similar linewidth of quaternary ^{14}N nuclei) of nitrogen nuclei (Mann, 1978a) compared to ^{13}C nuclei in trimethylamino groups provides the possibility for improved resolution of resonances. The availability of ^{13}C and ^{14}N nuclei as probes for amino groups in proteins also provides the researcher with choice. Moreover, the application of ^{14}N NMR spectroscopy

to the study of N^{α}- and N^{ε}-amino groups in protein should provide a convenient, cost-effective, quantitative and rapid method for observing the reaction of protein amino groups with iodomethane and will serve as a useful tool in protein structure-function studies.

The focus of the present study was to examine the feasibility of the employment of ^{14}N NMR spectroscopy as a general tool to characterize the trimethylamino derivatives (betaines) of several *in vacuo* methylated model proteins and of methylated amino acid (trimethylamino) standards. The latter were prepared *in vacuo* for the first time using the technique of Taralp and Kaplan (1997); Taralp (1997; Doctoral Thesis) has prepared the trimethylamino derivatives of amino acids in basic aqueous solutions from relatively large amounts of precursor amino acids (50 mg vs 2-5 mg). Exhaustive methylation was also employed for the synthesis of betaine amino acids as spiking standards (Greenstein and Winitz, 1961).

1.1 Rationale

The strategy of the present work consists of several key features. N^{α}- and N^{ε}-amino groups of protein are *in vacuo* methylated with iodomethane using the methodology of Taralp and Kaplan (1997). Methylated proteins are acid hydrolyzed to liberate the derivatized amino acid residues. It is important to ensure that derivatized groups remain intact following hydrolysis of the protein into the component amino acids. Methylation is a chemical modification that is stable towards the conditions used for conventional acid hydrolysis. The free N^{α}- and N^{ε}-trimethylated amino acids are analyzed by ^{14}N NMR spectroscopy and resonances are assigned by the aid of amino acid trimethyl-amino standard compounds. Moreover, by comparison to an internal standard

compound, tetramethylammonium chloride, it is possible to quantify the amount of each trimethylated amino acid derived from protein.

Sharp resonances are expected to result from the liberated amino acid trimethylamino derivatives (Zhang *et al.*, 1994) and these can be exploited for identification and quantification purposes. Taralp and Kaplan (1997) have shown that methylation of proteins with iodomethane targets other residues besides N^{α}- or N^{ϵ}-amino groups, namely histidine, cysteine, tyrosine and methionine. The nitrogen atom of the derivatized histidine residue, and side-chain nitrogen atom(s) of amino acids such as arginine, tryptophan and histidine (underivatized), does not give rise to any detectable signals. The electronic distribution surrounding these nitrogen atoms is very unsymmetrical, and their motions are relatively slower (cf. N^{α}- and N^{ϵ}-trimethylated amino acids), leading to extensive broadening of their quadrupolar ^{14}N NMR resonances.

Moreover, the electronic symmetry around the ^{14}N nucleus of the liberated amino acids that are not derivatized at the α-amino group is subject to change with pH, being of highest symmetry at low pH as the quaternary ammonium cationic moiety. The protonated nitrogen ($R\text{-}N^{+}H_3$) is nearly tetrahedral and leads to sharp resonances (Tzalmona and Loewenthal, 1974; Cohen *et al.*, 1975). When the pH of the sample is increased, the α-amino group of these amino acids is in the deprotonated form, which results in a more unsymmetrical electronic environment surrounding the quadrupolar ^{14}N nucleus and leads to line broadening of the ^{14}N resonance. Advantage of this line broadening as a result of deprotonation is taken. The analysis or resonance detection is simplified by performing the ^{14}N NMR spectroscopic analysis at high pD 11.4 (pH meter reading of 11), where the amino groups of liberated underivatized amino acids, which are

present in large excess following acid hydrolysis, become so broad that their signals are not observed as they disappear into the noise, leaving only the sharp resonances of the trimethylated amino acids derived from methylated protein. At high pD values, the resonance for an underivatized amino acid is a composite and weighted average of acidic and basic species, which increases the linewidth as a result of the loss of tetrahedral symmetry that accompanies the removal of the deuterium ion from nonquaternary alkylamino groups. In the case of the trimethylamino acids, the nitrogen atom is quaternized using nondissociable methyl groups, thus eliminating the possibility of deuterium ion exchange at all pD values, and the loss of tetrahedral symmetry at high pD values.

The focus of this investigation was to identify various amino acids of protein by the ^{14}N NMR chemical shift of their trimethylated amino groups. The success of the approach rests on the premise that the chemical shift of the various amino acid trimethylamino derivatives will each be disperse enough compared to the linewidth (approximately 3 Hz; Zhang *et al.*, 1994) to be identified from each other on the basis of the differences in their side-chain structure.

If structure-function studies are desired, the protein obviously cannot be hydrolyzed, and linewidths for *in vacuo* derivatized protein are expected to be comparable to that measured by Zhang *et al.* (1994) for calmodulin (23 Hz). ^{14}N NMR spectroscopy has been shown to be noninvasive for intact proteins and further information about the mobility of trimethylamino group (e.g., side-chain of trimethyl-amino lysine) can be obtained from ^{14}N relaxation measurements (Zhang *et al.*, 1994).

1.2 Background Information on Nitrogen NMR

The ^{14}N resonance shift between NH_4^+ and NO_3^- in the compound NH_4NO_3 was one of the first chemical shifts to be reported (Proctor and Yu, 1950). This observation opened the door to the examination of nuclei by NMR (Proctor and Yu, 1950). Both isotopes, ^{14}N and ^{15}N, are observable by NMR, and the properties of both nuclei have been reviewed (Mann, 1978a; Witanowski and Webb, 1972a; Mooney and Winson, 1969).

While ^{14}N and ^{15}N each has merit as an NMR active probe, there are drawbacks for either nucleus. With a natural abundance of 99.63 % and a gyromagnetic ratio of 7.224×10^{-2} rad^{-1} ($\nu = \gamma_N B_0 = 7.224 \times 10^{-2}$ rad^{-1} x 500 MHz = 36.12 MHz = the resonance frequency of ^{14}N nucleus on a 500 MHz spectrometer) relative to that of ^{1}H, ^{14}N has a relative sensitivity of 0.00101 (Witanowski and Webb, 1972a), and 5.69 with respect to naturally abundant ^{13}C nuclei (Witanowski and Webb, 1972a). Although ^{14}N signals are easily detected, this nucleus also possesses a relatively large quadrupole moment (I=1) of 7.1×10^{-2} (units: e x 10^{-24} cm^2) compared to deuterium (Witanowski and Webb, 1972a), and consequently, the ^{14}N signals are generally broad with linewidths ($\Delta\nu_{1/2}$) anywhere between 100 and 1000 Hz, depending on the electronic symmetry surrounding the nitrogen nucleus. Moreover, although the chemical shift range of nitrogen typically spans 1000 ppm, the chemical shift range of a given class of nitrogen compounds is usually on the order of 50 ppm, and as such, broad lines can prove problematic (unresolvable) in cases where two or more resonances are in close proximity (Mann, 1978a). In addition to impeding resolution, the rapid quadrupolar relaxation of the ^{14}N nucleus may remove information relating to the coupling of ^{14}N to other nuclei (Witanowski and Webb, 1972a), but with improvements in instrumentation, many of the

potential problems have been minimized or alleviated altogether (Witanowski and Webb, 1972a). As such, the ^{14}N nucleus has generally been used in chemical shift studies on natural abundance samples and in quadrupolar relaxation studies to obtain structural information (Witanowski *et al.*, 1981a) and information on molecular dynamics (Mann, 1978a).

In contrast, ^{15}N has a spin quantum number, $I = ½$, and yields sharp resonances ($\Delta\nu_{½} = 1$ Hz) and well-defined coupling constants (Witanowski and Webb, 1972a). It would be the isotope of choice for chemical shift measurements, J couplings and various structural studies (Witanowski *et al.*, 1981a), if not for its low natural abundance of 0.365 % and negative nuclear Overhauser enhancement due to its negative γ value (Witanowski and Webb, 1972a), making ^{15}N an extremely insensitive nucleus. Until recently, this has meant that ^{15}N studies were limited to neat low molecular weight compounds or isotopically enriched compounds, the latter route representing a costly and labourious undertaking (Witanowski and Webb, 1972a). Fortunately, the advent of wide-bore superconducting magnets and other techniques has somewhat narrowed this impasse (Mann, 1978a); nevertheless, ^{14}N and ^{15}N nuclei each have inherent advantages and disadvantages.

It follows, therefore, that in a structural study requiring sharp resonance peaks, the nucleus of choice should combine the sharp resonance line shape of ^{15}N with the high natural abundance, and hence convenience and sensitivity, of ^{14}N. Indeed, under special circumstances, a ^{14}N nucleus could offer all these features.

1.2.1 Quadrupolar Relaxation Mechanism for ^{14}N Nuclei with Asymmetrical Electronic Distribution Yields Broad Lines

Unlike ^{15}N nuclei, ^{14}N nuclear relaxation is dominated by the quadrupolar mechanism. All nuclei with spins greater than one-half ($I > ½$) have an electric quadrupole moment, an associated electric field gradient at the nuclear site, and a magnetic dipole moment. The quadrupole moment interacts with the inhomogeneous electric field or the electric field gradient around the nucleus to provide a torque on the nucleus, and this has an associated energy term. Electric field gradients originate from the electrical asymmetries in the molecule. In NMR, spin-lattice relaxation from the excited state to the ground state occurs when the magnetic moment of the nucleus interacts with magnetic fields fluctuating at its Larmor (or resonance) frequency (ω_0). In quadrupolar nuclei, the origin of these fluctuating magnetic fields is molecular tumbling of the electric field gradient interacting with the quadrupole magnetic moment of the nucleus (Reisse, 1982). Thus, nuclear spins of quadrupolar nuclei interact with magnetic fields to give rise to quadrupolar and Zeeman energy terms, respectively (Harris, 1983a). Electric field gradients vary in magnitude as the inverse cube of the distance from a charge and therefore, gradients which influence the magnetic properties (moment) of nuclei must originate from charge density variations that are near the nucleus (Akitt, 1972). Thus, the electric field gradient is influenced by the arrangement of neighbouring bonds and the electron density associated with each bond near the nucleus. For example, tetramethylammonium ion and benzyltrimethylammonium ion have similar electronic distributions and electric field gradients near the nitrogen center, as evidenced by the similar observed linewidths of 7 and 12 Hz, respectively (Akitt, 1972).

The expression for the longitudinal or nuclear spin-lattice relaxation time (T_{1N}=T_{1Q}) and the transverse or nuclear spin-spin (T_{2N}=T_{2Q}) relaxation time is given by:

$$\pi\Delta\nu_{1/2} = 1/T_{1Q} = 1/T_{2Q} = 3/8\ (1 + \eta^2/3)\ (eqeQ/h_{bar})^2\ \tau_c$$

where ω_0 is the Larmor frequency for the ^{14}N nucleus in hertz, τ_c is the effective correlation time in seconds for reorientation at the site of the functional group in the molecule containing the ^{14}N nucleus, $\Delta\nu_{1/2}$ is the resonance width at half height in hertz (the linewidth); T_Q is the quadrupolar relaxation time, ($eqeQ/h_{bar}$ or $eq_{zz}eQ/h_{bar}$) is the nuclear quadrupolar coupling constant (NQCC) in hertz, eq is the electric field gradient, eQ is the nuclear quadrupole moment, and η describes the deviation of the electric field gradient from axial symmetry [$\eta = (q_{xx} - q_{yy})/q_{zz}$, where q_{xx}, q_{yy} and q_{zz} are the principal components of the electric field gradient tensor] (Witanowski and Webb, 1972b; Gerothanassis *et al.*, 1987).

The expression for the relaxation mechanism given above incorporates the extreme narrowing approximation where $\tau_c << 1/\omega_0$. The latter applies to small or medium-sized molecules in nonviscous liquids in the absence of line-broadening contributions due to chemical exchange phenomena (Harris, 1983a). When ^{14}N nuclear relaxation is efficient, the linewidths are broad. The linewidth ($\Delta\nu_{1/2}$) is a function of the nuclear quadrupole moment, electric field gradient, and correlation time. The rotational contribution to the linewidth (i.e., τ_c) is normally minimal since ^{14}N nuclear relaxation is usually dominated by the quadrupolar coupling constant term (eqeQ/h) (Drago, 1992a). This is not the situation with symmetrical quaternary nitrogens (e.g., tetraalkylammonium salts), where narrow lines are observed.

1.2.2 Dipolar Relaxation Mechanism for ^{14}N Nuclei with Symmetrical Electronic Distribution Yields Narrow Lines

Typically, ^{14}N relaxation times range from 10 to 100 ms and this explains why the linewidth, which is inversely related to the spin-spin relaxation time, T_{2Q}, is often quite broad (Mann, 1978b). However, the magnitude of the QCC is dependent on the magnitude of the electric field gradient at the nucleus (Witanowski and Webb, 1972b). In short, an electric field gradient is necessary in order to bring about quadrupolar relaxation (Witanowski *et al.*, 1981b), so that in cases of very high symmetry about the nitrogen nucleus, an electric field gradient may be very small or absent and would result in relaxation by other processes (Witanowski *et al.*, 1981b), which may span much longer times, as in the case of the $(CH_3)_4N^+$ ion (Mann, 1978b). Since the linewidth at half height ($\Delta\nu_{1/2}$) is inversely proportional to the spin-spin relaxation time T_{2N}, it is expected that ^{14}N nuclei of high electronic symmetry would have sharp linewidths approaching those of ^{15}N nuclei. This has been shown to be the case for many trigonal planar (reference NO_3^-) (Witanowski and Webb, 1972c) and tetrahedral (reference R_4N^+) (Witanowski and Webb, 1972d) nitrogen compounds; the latter compounds giving particularly narrow lines. Ion pairing of tetraalkylammonium salts results in the distortion of the electronic symmetry about the ^{14}N nucleus, and consequently, the resonance line is broadened (Harris, 1983a). This demonstrates the ease by which the electronic symmetry can be distorted. This sensitivity to environment was the basis for the experiments conducted on ^{14}N-nitrate binding to erythrocyte protein AE1, which were carried out by Galanter and Labotka (1991).

Other relaxation processes come into play in the absence of a strong electric field gradient about a quadrupolar nucleus, where quadrupolar relaxation is not efficient and

has a minimal contribution to the overall relaxation process. These other relaxation processes are less efficient (i.e., longer relaxation times) and become the principle factors that define the resonance linewidth (Witanowski *et al.*, 1977b). As with spin ½ nuclei, the other possible mechanism of relaxation for a quadrupolar nucleus results from dipole-dipole interactions (i.e., of two nuclear magnetic moments), which are modulated by molecular motions (tumbling, or translational diffusion). In NMR, nuclear energy transitions resulting in spin-lattice relaxation occur when the magnetic dipole moment of the nucleus interacts with magnetic fields that fluctuate at its Larmor (or resonance) frequency (ω_0) (Breitmaier and Voelter, 1987b). In the absence of quadrupolar relaxation, the magnetic dipole of the quadrupolar ^{14}N nucleus interacts with local fluctuating magnetic fields, and this interaction has an associated energy term. The origin of the fluctuating magnetic fields is molecular tumbling of the magnetic field from a neighbouring nuclear magnetic moment, or from unpaired electrons or electronic clouds that are around the quadrupolar ^{14}N nuclear magnetic moment, which gives rise to the chemical shielding (i.e., around the ^{14}N nucleus which is being observed). Thus, in the absence of a strong electric field gradient about a quadrupolar nucleus, nuclear spins of quadrupolar nuclei interact with fluctuating magnetic fields to give rise to dipolar and Zeeman energy terms, respectively (Harris, 1983a). Dipole-dipole interactions can be intra- or intermolecular, or homo- or heteronuclear. For example, the relaxation of $ND_4^+Cl^-$ below 220 K was dominated by intramolecular N-D dipolar coupling interactions. In aqueous solution, intermolecular dipole-dipole interactions relax nitrogen nuclei of ammonium salts by the reorientation of water dipoles and their associated local

magnetic fields (Witanowski *et al.*, 1977b). Both inter- and intramolecular dipolar mechanisms can contribute to the relaxation of tetraalkylammonium salts.

1.2.3 Correlation Times and Line Broadening in Proteins

Energy transfers resulting in relaxation will occur if the position vectors that determine instantaneous magnetic coupling of a nucleus with its surroundings are a function of time. The variations between these "position" functions over short intervals of time are characterized by the auto-correlation function, $k(\tau)$, which in turn, is characterized by the correlation time, τ_c. The auto-correlation function, $k(\tau)$, models the exponential decay of the position of magnetic vectors with time due to Brownian motion, as defined by a statistical treatment of the loss of correlation by the ensemble of nuclear spins (Harris, 1983). When the auto-correlation function is Fourier transformed, it is converted to a frequency domain function known as the spectral density function, $J(\omega)$, which describes the density of single quantum transitions possible at a particular frequency.

The auto-correlation and spectral density functions have been used to calculate Zeeman level transition probabilities of nuclei and to determine the dependence of various relaxation processes on correlation time (τ_c), and hence, on molecular tumbling, which is at the origin of almost all relaxation mechanisms (Reisse, 1982; Drago, 1992b). All relaxation processes have a dependency on correlation time. The efficiency of longitudinal relaxation (T_1) is affected by the external magnetic field strength (Reisse, 1982; Harris, 1983b). The relaxation is more favoured if $J(\omega)$ is large and the relaxation occurs for motions at the Larmor frequency, where $J(\omega_o)$ is proportional to $\tau_c/(1+ \omega_o^2\tau_c^2)$ and to the relaxation rate T_1^{-1}. T_1 is most efficient when $\tau_c = 1/ \omega_o$.

A sharp peak usually arises from a small molecule in a non-viscous solvent. In the "extreme narrowing" condition, where $\omega_o^2\tau_c^2<<1$, the longitudinal relaxation time, T_1, is equal generally to the transverse relaxation time, T_2, and linewidths are narrow ($\Delta\nu_{1/2}=1/\pi T_2$), where T_1^{-1} is proportional to τ_c. Proteins are macromolecules and have slower molecular motions compared to small molecules, and hence have different relaxation times ($T_2 \neq T_1$), relaxation rates ($T_2^{-1} \neq T_1^{-1}$), and spectral density functions (at the "non-extreme narrowing" condition $\omega_o^2\tau_c^2>>1$, T_2^{-1} is proportional to τ_c; while T_1^{-1} is proportional to τ_c^{-1}). The transverse relaxation (T_2), which affects linewidth, becomes increasingly efficient as correlation times (τ_c) (and hence, molecular weights) increase. Hence, a ^{13}C atom in a macromolecular protein will have a broader resonance than that in a small molecule in a non-viscous solvent ($\Delta\nu_{1/2}=1/\pi T_2$ and $\Delta\nu_{1/2}$ is proportional to τ_c). Zhang *et al.* (1994) have shown that the linewidths were narrow for free trimethylamino lysine (3 Hz). In this case, a rapid average tumbling rate promoted efficient relaxation. In contrast, the trimethylamino lysyl residue in cytochrome c or calmodulin gave linewidths of 23 Hz due to the much larger and slower tumbling macromolecules, which had longer correlation times.

2. Materials and Methods

2.1 Proteins

Insulin (bovine; Lot), cytochrome c (horse heart; Lot 27H-7065), ribonuclease A (bovine; Lot), lysozyme (chicken egg-white, grade I; Lot 89F8275), α-chymotrypsin (bovine; Lot 91H7195), and chymotrypsinogen A (bovine pancreas; Lot 39F 8095) were purchased from Sigma Chemical Company.

2.2 Chemicals and Solvents

^{12}C-Iodomethane was purchased from Sigma-Aldrich Chemical Company. L-Amino acids (Ala, Arg, Asp, Cys-SO_3^-, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val), cystine, and poly-L-Lys·HBr were purchased from Sigma Chemical Company. All other chemicals, reagents, and solvents were high-purity preparations obtained from commercial sources.

2.3 *In Vacuo* Methylation of the Amino Acids: Preparation of ^{12}C-N-Trimethylated Amino Acid Standards

Amino acid or Poly-L-Lys HBr (5 mg; 5 mL of a 1 mg/mL stock solution prepared in dH_2O) solution adjusted to pH 10.5 was lyophilized (LpH 10.5) in a Pyrex screw-cap tube. Methylation was carried out *in vacuo* at 100-105°C for 4 days with iodomethane (20 µL for each amino acid or 30 µL for Poly-L-Lys HBr) according to the procedure described by Taralp and Kaplan (1997) for proteins. L-Tyr (5 mg each sample) was also reacted at LpH 11 and at LpH 9.3. In some cases (L-His, L-Trp, L-Cys–SO_3^-, L-Glu, and L-Asp), 20 mg amino acid were dissolved in 1 mL dH_2O and sufficient base to set the pH of the solution to 10.5 and subsequently lyophilized. In such cases, 60 or 80 µL of iodomethane were used for the *in vacuo* methylation. Methylated poly-L-Lys was acid hydrolyzed in 6N HCl at 110°C for 24 h.

2.4 Exhaustive Methylation of the Amino Acids: Preparation of ^{12}C-N-Trimethylated Amino Acids (Betaines) Spiking Standards

Synthesis of the betaines of the various amino acids was carried out according to the protocol for exhaustive methylation, as presented in Greenstein and Winitz (1961). Precursor amino acid (200 mg) was dissolved in 5 mL anhydrous methanol and enough methanolic NaOH solution (1.6 mL 10% NaOH in anhydrous MeOH) was added to

produce a weakly alkaline reaction mixture. Iodomethane (5 mL) was added and the mixture was refluxed for 8 h, with small amounts of NaOH solution being added periodically to maintain a faintly alkaline reaction mixture. The solution was cooled and evapourated to dryness. Some betaine amino acids were evapourated to dryness, while others were oily residues that were crystallized quickly on ice and recrystallized from cold water. The products were the methyl esters of the trimethylamino acids in the iodide form. Warming in 1% aqueous NaOH solution for 1 h saponified the products and the amino acid betaines were retrieved after the samples were evapourated.

Samples were analyzed by thin-layer chromatography (TLC) using basic alumina TLC plates purchased from Sigma. Development was carried out to a height of 15 cm from the point of application. Separation of the amino acid betaines (amino acids with a quaternary ammonium group) was made with chloroform-methanol-ammonia (6:3:1) as developing solvent (McLean and Jewers, 1972). The separated materials were detected using a 2% solution of iodine in methanol, with an overspray of saturated sodium nitrite solution to intensify spots. The observed R_f values for the various products were in the range 0.37-0.76 (0.39-0.46, 0.6-0.76 for TLC 1; 0.37-0.5, 0.63 for TLC 2; see Appendix 4-1, Figures 4-1(a)A and 4-1(b)A), as expected with the conditions employed for such compounds (McLean and Jewers, 1972). Compounds with R_f values >0.5 were likely due to the product methyl esters not being completely saponified. The afore-mentioned developing solvent system has been shown previously to be successful in separating methyl ester betaine (i.e., trimethyl Gly methyl ester) and betaine (trimethyl Gly with a free carboxylic acid group) on basic alumina TLC plates (McLean and Jewers, 1972). Underivatized amino acids migrated very little (R_f = 0-0.09). ^{14}N NMR spectra

yielded more distinct resonances because of the larger quantity of amino acid betaines synthesized (e.g., 200 mg) and analyzed cf. those for *in vacuo* methylation reactions on lyophilized amino acid (5 mg; LpH 11). The chemical shifts of the betaine amino acids prepared by the exhaustive methylation method were similar to those obtained for the respective products as prepared via *in vacuo* chemical modification (see Table 4-1) and linewidth measurements were made easier for the amino acid betaine samples prepared by the former protocol.

If further purification of the products is required, cation-exchange chromatography can be employed using a Dowex 50-X8 (200-400 mesh; hydrogen form) column (1x108 cm) (Delange, Glazer, and Smith, 1969). Removal of acidic and neutral amino acids can be accomplished by washing with 500 mL 1 M pyridine, and the product can be eluted with 1 M pyridine-acetic acid at pH 5.2. Detection of contaminating amino acids can be made by spotting small aliquots of sample from each tube on paper and treated with ninhydrin/Cd acetate reagent. Also, TLCs can be run as described above on samples before and after the cation-exchange chromatography. [Other useful literature: Fried and Sherma, 1994; Jork *et al.*, 1990; Muller and Eckert, 1989; Reuvers *et al.*, 1989; Tyihak *et al.*, 1974; Snyder *et al.*, 1986.]

2.5 *In Vacuo* Methylation of Model Proteins with ^{12}C-Iodomethane

Protein (typically 10-30 mg except where stated otherwise) was reacted *in vacuo* with 25 µL ^{12}C-iodomethane at LpH 2.3 (insulin only), 7, 7.5 (insulin only), 9.2 or 10 (from a lyophilization volume or Lvol of 1 mL; Lvol 20 mL for insulin), according to the procedure of Taralp and Kaplan (1997). Samples at LpH 10 were lyophilized from aqueous solution containing 40 mM sodium metaborate buffer, pH 10 (2 mM for insulin),

while those at LpH 7 or 7.5 were prepared from aqueous solution containing 40 mM sodium phosphate buffer, pH 7 or 7.5 (2 mM for insulin), adjusted to the appropriate pH with acid or base as required. Methylated protein samples were acid hydrolyzed (Darbre, 1986) and subsequently analyzed by ^{14}N NMR spectroscopy. α-Chymotrypsin and chymotrypsinogen A samples were acid hydrolyzed (Darbre, 1986) and subsequently performic acid oxidized (Hirs, 1956) prior to NMR analysis.

2.6 Acid Hydrolysis of ^{12}C-Methylated Proteins

Methylated protein was transferred into a new Pyrex hydrolysis tube and 1 mL 6N HCl was added. The resultant suspension was frozen, degassed and sealed under vacuum by flame. Hydrolysis was carried out *in vacuo* at 110°C for 24 h (Darbre, 1986). Sample tubes were then cracked open and the acid was removed under vacuum in a desiccator containing sodium hydroxide pellets and P_2O_5. The subsequent addition of 1 mL distilled water to each sample and re-freeze-drying removed residual acid. Methylated proteins were hydrolyzed because the correlation time for a protein, τ_c, is slow or long *vs* the linewidth at half height, $\Delta\nu_{1/2}$, which is large. [For a trimethylamino group on a protein, $\Delta\nu_{1/2}$ is ≈ 30 Hz; for trimethylamino acids, $\Delta\nu_{1/2}$ is ≈ 3 Hz (Zhang *et al.*, 1994).]

2.7 Performic Acid Oxidation of Methylated α-Chymotrypsin and Chymotrypsinogen Hydrolysates

Performic acid oxidation of chymotrypsin and chymotrypsinogen A samples was carried out according to the procedure of Hirs (1956; referenced in Bailey, 1964). Performic acid solution was prepared by incubating a mixture of 0.5 mL 30% hydrogen peroxide and 9.5 mL of 98-100% formic acid at room temperature for 2 h in a screw cap hydrolysis tube. The mixture was divided equally and is sufficient to react two samples

of protein up to 30 mg per sample. Normally, performic acid oxidation is carried out on protein, but it was performed here on hydrolyzed protein, which should not pose a problem. Each protein hydrolysate sample was dissolved in 2.5 mL of 99% formic acid and 1.0 mL anhydrous methanol in a Pyrex hydrolysis tube; the latter is added to prevent freezing of the solution later on in the protocol. The three hydrolysis tubes containing the two protein hydrolysate solutions and the performic acid solution were placed in an ice-NaCl bath at –10°C and cooled for 30 min, at which time the performic acid solution was split and added to each sample. After a reaction time of 2.5 h, the content of each tube was rinsed with 50 mL of ice water and transferred into a 1 L round bottom flask containing 300 mL of water at 0°C. The samples were shelled with liquid nitrogen and lyophilized for 3 days. Samples were re-dissolved in 10 mL water and re-lyophilized (in a scintillation vial or pointy bottom flask for facile sample retrieval) to remove traces of performic acid. Samples were fluffy light beige lyophilized powders.

2.8 ^{14}N NMR Spectroscopic Analysis of ^{12}C-Methylated Amino Acid Standards

Each derivatized amino acid sample was dissolved in 400 or 900 μL D_2O and transferred to an eppendorf tube. Tetramethylammonium chloride (TMACl) (30 μL of 10 mg/100 μL stock solution = 3 mg) was added and used as an internal standard, which was referenced at 0 ppm. TMACl is 333.5 ppm upfield of neat nitromethane, which is the most accepted standard (Witanowski and Webb, 1972d). Sample "pH" (pD) was adjusted to 11.4 with 1N and/or 5N NaOH and 1N HCl (if necessary) with vortex mixing. Centrifugation using a microcentrifuge at 15 000 g for 2 min separated insolubles from solubles if required. The supernatant was transferred to a fresh eppendorf tube, in which it was stored at -20°C until it was analyzed by ^{14}N NMR spectroscopy. Methylated

Poly-L-Lys HBr was first hydrolyzed (6N HCl, 110°C, 28 hr; dried in a desiccator under vacuum in the presence of P_2O_5 and NaOH pellets; +1 mL dH_2O and dried in the same manner to remove residual acid) and then prepared as above using 400 µL D_2O. Typically, 1024 scans were acquired for the analysis of standard compounds. A 1 Hz exponential linebroadening function was applied to the free induction decay prior to Fourier transformation. A 0.8 s acquisition time, 90° pulse (36 µs) and 0 s relaxation delay were used. Observed linewidths were measured directly from the spectra (processed using WIN-NMR software for each relevant peak, plotted at ±10 Hz (1.02 Hz/cm) to ensure accuracy in the measurement). The spectral window was 12 000 Hz. Proton decoupled spectra were obtained.

2.9 ^{14}N NMR Spectroscopic Analysis of ^{12}C-Methylated Proteins

^{14}N NMR spectra were obtained using a Bruker spectrometer operating at 9.4 Tesla (^{1}H, 400.13 MHz; ^{14}N, 28.9 MHz) using a probe that accommodates the ^{14}N-resonance condition (28.9 MHz). Methylated protein hydrolysates were dissolved in 700 µL D_2O and transferred to an eppendorf tube. Tetramethylammonium chloride (TMACl) (30 µL of 10 mg/100 µL stock solution = 3 mg) was added and used as an internal standard, which was referenced at 0 ppm. Sufficient 1N and 5N NaOH and 1N HCl (if necessary) was added with vortex mixing to give a pH meter reading of 11 (pD 11.4). Samples were run by ^{14}N NMR spectroscopy at pD 11.4, where trimethylated amino acids give rise to sharp peaks and other derivatives or underivatized sample peaks are very broad, and generally are unresolved in the baseline (Zhang *et al.*, 1994). Samples were stored at -20°C if necessary until they were analyzed by ^{14}N NMR spectroscopy. Typically, 1024 scans were acquired for the analysis of methylated protein

hydrolysates; however, in some cases, overnight scans (e.g., 40 000 scans) were required to observe expected resonances. A 1 Hz exponential linebroadening function was applied to the free induction decay prior to Fourier transformation. A 0.8 s acquisition time, 90° pulse (36 μs) and 0 s relaxation delay were used. Observed linewidths were measured directly from the spectra (processed for each relevant peak plotted at ±10 Hz (1.02 Hz/cm) to ensure accuracy in the measurement). The spectral window was 12 000 Hz. Proton decoupled spectra were obtained.

3. Results

3.1 ^{14}N NMR Spectroscopy of N^{α}- or N^{ε}-Trimethylamino Derivatives of the Amino Acids

The linewidths for the amino acid trimethylamino derivatives with small alkyl side-chain residues e.g., Ala, Gly, Leu, and Lys (ε), were observed to be the narrowest (see Table 4-1). Derivatives with bulkier alkyl side-chains such as Ile and Val gave larger linewidths, while even broader resonances were observed for amino acid trimethylamino derivatives consisting of more complicated side-chain functional groups, such as those for Arg, Glu, His, Met, Phe, Ser, Thr, Trp, and Tyr. Hexamethylamino derivative of cystine gave two very broad lines.

3.2 ^{14}N NMR Spectroscopy of Trimethylated N^{α}- and N^{ε}-Amino Groups in Protein Hydrolysates

Although amino groups were not the only functional groups to be methylated with iodomethane, their trimethylated derivatives were the only ones observable by ^{14}N NMR sprectroscopy as sharp lines at high pD values. This was in contrast to the results obtained by ^{13}C NMR spectroscopy, where aspartyl, glutamyl, tyrosyl, methionyl, cystyl

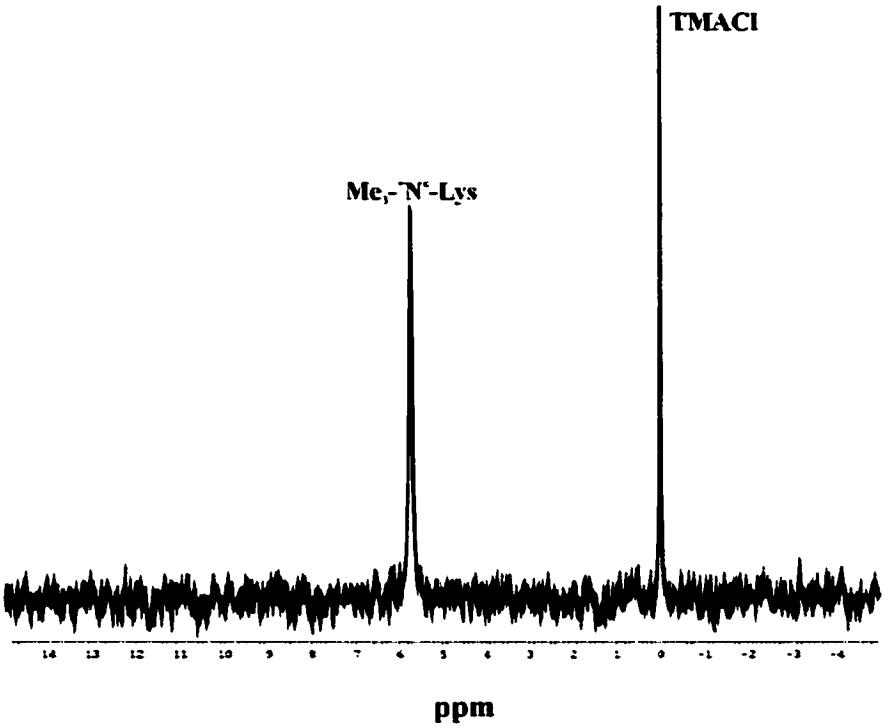


Figure 4-1. Expanded ^{14}N NMR spectrum (NT = 1024 scans) of ^{12}C-methylated cytochrome c (20 mg, LpH 10) hydrolysate (reacted *in vacuo* with ^{13}C-iodomethane at 75°C for 24 h, followed by hydrolysis *in vacuo* with 6 N HCl).

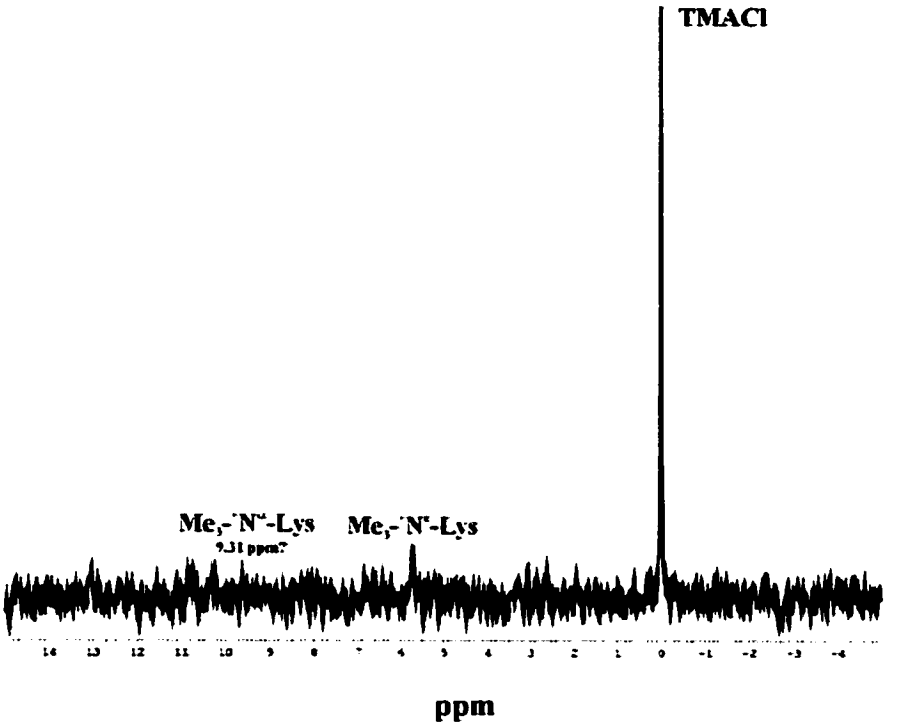


Figure 4-2. Expanded ^{14}N NMR spectrum (NT = 1024 scans) of ^{12}C-methylated ribonuclease A (20 mg, LpH 9.2) hydrolysate (reacted *in vacuo* with ^{13}C-iodomethane at 75°C for 24 h, followed by hydrolysis *in vacuo* with 6 N HCl).

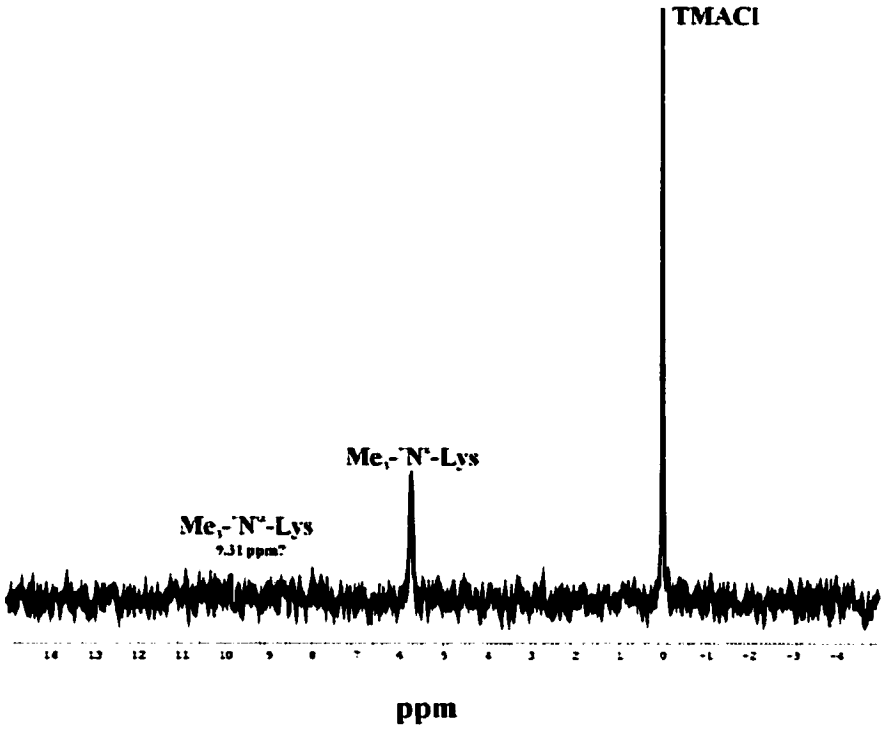


Figure 4-3. Expanded ^{14}N NMR spectrum (NT = 1024 scans) of ^{12}C-methylated lysozyme (20 mg, LpH 9.2) hydrolysate (reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h, followed by hydrolysis *in vacuo* with 6 N HCl).

and histidyl derivatives also are observed (Taralp and Kaplan, 1997; Chapters 1, 2, and 3). Each of the ^{14}N NMR spectra of the protein hydrolysates studied revealed a peak for the superimposed resonances of N^{ϵ}-trimethylamino lysine derivatives (Table 4-2 and all Figures). As expected, there was no other resonance observed than that for N^{ϵ}-trimethylamino lysine in the spectrum for the hydrolysate of cytochrome c, since the N-terminal glycyl residue is acetylated (Dayhoff, 1972) (Figure 4-1). The spectra for ribonuclease A (Figure 4-2) and lysozyme (Figure 4-3) did not reveal a resonance for N^{α}-trimethylamino lysine in addition to the observed resonance for N^{ϵ}-trimethylamino lysine. This was surprising since a resonance for N^{α}-trimethylamino lysyl residue was observed in the ^{13}C NMR spectra of ^{13}C-methylated derivatives of these proteins (Chapters 2 and 3). These results illustrate the sensitivity problem, as more protein was required to observe the trimethylamino resonances derived from N-terminal residues; moreover, more scans were required. For chymotrypsinogen A (Figure 4-4), a resonance for the N^{α}-trimethylamino half-cystine (Cya or cysteic acid after performic acid oxidation and acid hydrolysis) derived from the N-terminal Cys residue (half-cystine) was not observed.

Bovine insulin contains two N-termini, and hence, two N^{α}-amino groups, namely that of glycine at position 1 of the A peptide chain and phenylalanine at position 1 of the B peptide chain, and one N^{ϵ}-amino group of lysine at position 29 of the primary amino acid sequence of the B peptide chain of insulin (Dayhoff, 1972). Three sharp resonances were observed in the ^{14}N NMR spectrum of the methylated insulin hydrolysate reacted at LpH 10 (Figures 4-5a). The resonance at zero ppm corresponded to the chemical shift of the reference, tetramethylammonium chloride, which has a highly symmetrical electronic

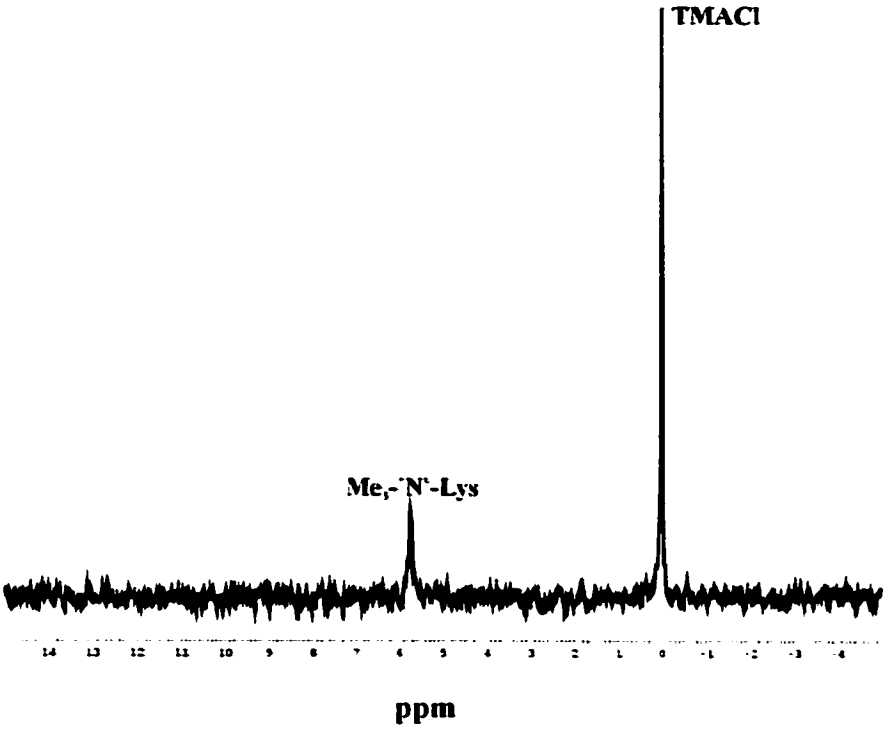


Figure 4-4. Expanded ^{14}N NMR spectrum (NT = 1024 scans) of ^{12}C-methylated chymotrypsinogen A (20 mg, LpH 10) performic acid oxidized hydrolysate (reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h, followed by hydrolysis *in vacuo* with 6 N HCl).

environment about the ^{14}N nucleus and was observed to be the sharpest with a linewidth at half-height of 0.6 Hz. The other two resonances at 5.77 and 4.33 ppm were attributed to the N^{ε}-trimethylamino lysyl and N^{α}-trimethylglycine, respectively, by comparison with the resonances of the respective trimethylated amino acid standards. The linewidth at half-height of the resonance at 5.77 ppm was 1.9 Hz, while that at 4.33 ppm was 1.7 Hz (taking the line broadening into consideration), the trend being consistent with that of the respective derivatized trimethylated amino acid standards (Table 4-2). However, the trimethylamino group of the N-terminal phenylalanyl residue of insulin was not observed due to the low number of scans taken (1024 scans). The standard for trimethylamino phenylalanine produced a slightly broad resonance peak. This, coupled to a sensitivity problem due to the small amount of derivatized phenylalanine in the sample as a result of incomplete hydrolysis of the bulky N-terminal trimethylamino Phe-Val dipeptide (see Chapter 2) resulted in the undetectable trimethylated phenylalanine derived from insulin. Figure 4-5b shows the ^{14}N NMR spectrum for a methylated insulin hydrolysate sample reacted at LpH 9.2 that was run overnight for 45 000 scans, which improved the S/N ratio and revealed a broad resonance at 10.24 ppm for the trimethylamino group of the phenylalanine derived from insulin. The correlation between the latter broad peak and the trimethylamino phenylalanine standard was not as good as was that between the other resonances and their corresponding trimethylamino amino acid standards (Tables 4-1 and 4-2).

Figure 4-5c shows the ^{14}N NMR spectrum (2048 scans) observed for methylated insulin hydrolysate sample reacted at LpH 7.5. Besides the reference peak, one sharp resonance was observed at 4.65 ppm having a linewidth at half-height of 3.2 Hz, and

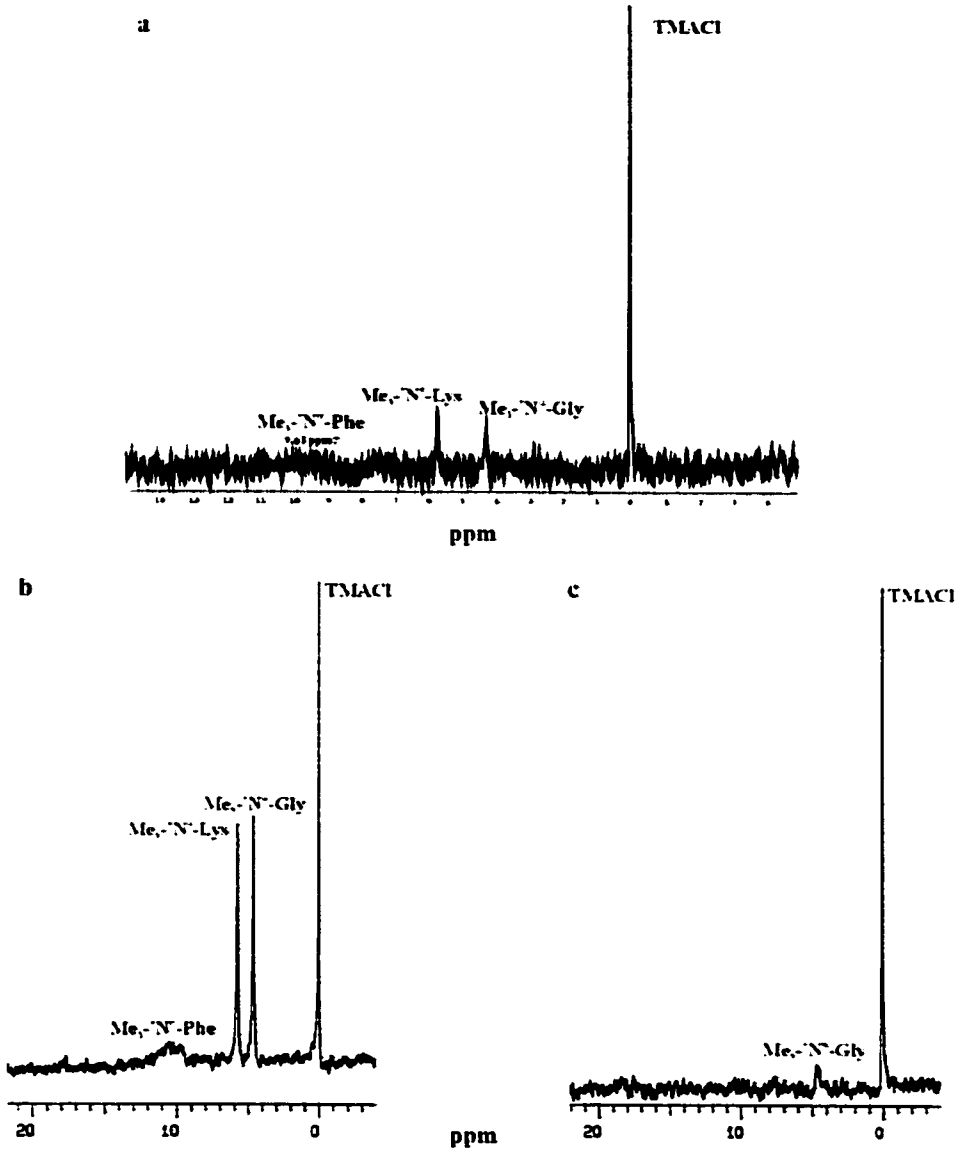


Figure 4-5. Expanded ^{14}N NMR spectra of ^{12}C-methylated insulin (20 mg) hydrolysate (reacted *in vacuo* with ^{13}C-iodomethane at 75°C, 24 h, followed by hydrolysis *in vacuo* with 6 N HCl) at (a) LpH 10 (NT = 1024 scans; Bruker 400 MHz spectrometer), (b) LpH 9.2 (NT = 45000 scans; Varian 300 MHz spectrometer), and (c) LpH 7.5 (NT = 2048 scans; Varian 300 MHz spectrometer).

corresponded to N^{α}-trimethylamino glycine. A resonance for N^{ϵ}-trimethylamino lysine was not observed. This was expected, since the pK_a of the N^{ϵ}-amino group of the lysyl residue of a protein is between 10 and 11 (Creighton, 1993a). A peak for N^{α}-trimethyl-amino phenylalanine was not observed for similar reasons as given above.

The hydrolysate of α-chymotrypsin (Figure 4-6) revealed three resonances in addition to that for the N^{ϵ}-trimethylamino lysine, which corresponded to the N^{α}-trimethylamino groups of isoleucine, alanine and glycine at 11.29, 9.11, and 4.58 ppm, respectively. Glycine is not an N-terminal residue of α-chymotrypsin and the observed peak that appears to correspond with the standard for N^{α}-trimethylamino glycine is likely the result of auto-proteolysis or contamination. Extra resonances corresponding to trimethylamino groups of other amino acid residues were observed in the ^{13}C-NMR spectra of ^{13}C-methylated α-chymotrypsin (see Chapters 2 and 3). As with chymotrypsinogen A, a resonance for the N-terminal half-cystine (Cya or cysteic acid after performic acid oxidation and acid hydrolysis) was not observed. Since the spectrum for the cystine standard (or Cya (cysteic acid) standard) gave two broad resonances at 14.75 ppm (12.1 Hz linewidth) and 9.89 ppm (9.7 Hz linewidth) (or for trimethylamino Cya standard, at 9.10 ppm, 13.7 Hz linewidth), it is likely that the resonance peak corresponding to it in the protein sample was too broad to be observed.

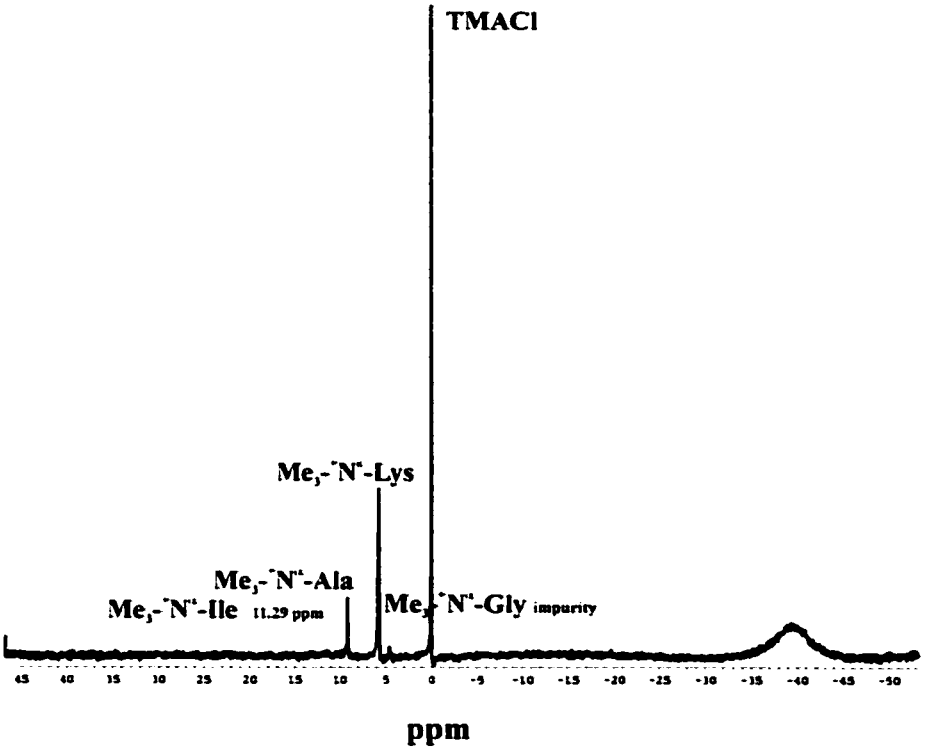


Figure 4-6. Expanded ^{14}N NMR spectrum (NT = 40 000 scans) of ^{12}C-methylated α-chymotrypsin (240 mg, LpH 10) performic acid oxidized hydrolysate (reacted in vacuo with ^{13}C-iodomethane at 75°C for 24 h, followed by hydrolysis in vacuo with 6 N HCl).

Table 4-1 Chemical Shifts of N^{α}- or N^{ε}-Trimethylamino Derivatives of the Amino Acids

Amino Acid	^{14}N Resonance† (ppm) [for *in vacuo* or (exhaustive) methylated samples]	Linewidth, $\Delta\nu_{1/2}$ (Hz) [for *in vacuo* or (exhaustive) methylated samples]
Ala	9.15	1.9
Arg	9.40 (9.39)	3.4 (7.9)
Asp	NO (8.86)	NO (7.2)
Cys	11.02	
Cya (cysteic acid)	9.10	13.7
Cystine	14.75, 9.91	12.1, 9.7
Glu	9.17 (9.16)	6.3 (8.2)
betaine	4.52	1.6
Gly	4.62	1.6
His	9.59 (9.26)	8.5 (21.1)
Ile	11.29 (11.24)	2.9 (5.9)
Leu	9.01 (8.94)	3.4 (5.4)
Lys (α)	9.27	3.5
Lys (ε)	5.74	3.6
Poly-L-Lys (ε) hydrolysate	5.74	2.0
Lys (α, ε)	9.31 (9.31), 5.75 (5.79)	3.5 (9.3), 3.6 (5.1)
Met	8.58 (8.51)	3.1 (9.7)
Phe	9.73 (9.58)	7.9 (9.3)
Pro	26.75 (26.55)	3.2 (7.0)
Ser	9.03 (8.92)	5.2 (8.5)
Thr	9.14	5.1
Trp	9.71 (9.69)	4.2 (6.7)
Tyr	9.58 (9.61)	9.3 (10.1)
Val	11.43 (11.40)	2.7 (5.6)

† Chemical shift referenced to internal tetramethylammonium chloride, at 0.00 ppm with a linewidth at half-height, $\Delta\nu_{1/2}$, of 0.6 Hz. After acid hydrolysis of the methylated protein, the N-terminal amino acid(s) is identified from a table of the chemical shifts of the N^{α}-trimethylamino derivatives of the amino acids. NO = not observed. **Confirmation of Identity:** Spike the sample with the standard trimethylated amino acid tentatively identified as an amino-terminus and observe whether its resonance corresponds exactly with that of the standard.

Table 4-2 Chemical Shifts of N^{α}- or N^{ε}-Trimethylamino Derivatives of Amino Acids Derived from Protein Hydrolysates

Trimethylamino acid standard	Standard ^{14}N Resonance† (ppm)	Linewidth $\Delta\nu_{1/2}$ (Hz)	^{14}N Resonance† (ppm)	Linewidth $\Delta\nu_{1/2}$ (Hz)
			Insulin LpH 10[a]	
$(Me)_4N^+Cl^-$	0	0.6	0	0.9
Lys (ε)	5.74	2.0	5.74	2.7
Gly	4.62	1.6	4.28	2.9
Phe	9.73	7.9	NO	-
			Insulin LpH 7.5[b]	
$(Me)_4N^+Cl^-$	0	0.6	0	0.8
Lys (ε)	5.74	2.0	NO	-
Gly	4.62	1.6	4.75	2.0
Phe	9.73	7.9	NO	-
			Insulin LpH 9.2[c]	
$(Me)_4N^+Cl^-$	0	0.6	0	0.9
Lys (ε)	5.74	2.0	5.73	2.8
Gly	4.62	1.6	4.60	3.0
Phe	9.73	7.9	(10.24)	-
			Cytochrome c LpH 10[d]	
$(Me)_4N^+Cl^-$	0	0.6	0	0.7
Lys (ε)	5.74	2.0	5.74	2.4
			Ribonuclease A LpH 9.2[e]	
$(Me)_4N^+Cl^-$	0	0.6	0	0.8
Lys (ε)	5.74	2.0	5.72	3.0
Lys (α)	9.27	3.5	NO	-
			Lysozyme LpH 9.2[f]	
$(Me)_4N^+Cl^-$	0	0.6	0	0.8
Lys (ε)	5.74	2.0	5.74	2.6
Lys (α)	9.27	3.5	NO	-
			Chymotrypsin LpH 10[g]	
$(Me)_4N^+Cl^-$	0	0.6	0	0.8
Lys (ε)	5.74	2.0	5.75	2.6
Ala	9.15	1.9	9.12	2.1
Cya	9.10	13.7	NO	-
Ile	11.29	2.9	11.29	2.0
Gly	4.62	1.6	4.57	2.7
			Chymotrypsin LpH 7[h]	
$(Me)_4N^+Cl^-$	0	0.6	0	0.9
Lys (ε)	5.74	2.0	5.73	3.5
Ala	9.15	1.9	9.16	5.1
Cya	9.10	13.7	NO	-
Ile	11.29	2.9	NO	-
Gly	4.62	1.6	4.59	2.7

			Chymotrypsinogen LpH10[i]	
$(Me)_4N^+Cl^-$	0	0.6	0	0.9
Lys (ε)	5.74	2.0	5.75	2.7
Cya	9.10	13.7	NO	-
			Chymotrypsinogen LpH7[j]	
$(Me)_4N^+Cl^-$	0	0.6	0	2.6
Lys (ε)	5.74	2.0	5.75	4.0
Cya	9.10	13.7	NO	-

NO = not observed.

4. Discussion

^{14}N NMR spectroscopy was employed as a tool to study the reaction of N^{α}- and N^{ε}-trimethylamino groups in methylated proteins and amino acid trimethylamino derivatives. In the majority of cases, the chemical shift resonances of trimethylamino acids derived from methylated protein hydrolysates correlated very well with amino acid trimethylamino standard compounds representing the expected reactable functional groups (Tables 4-1, 4-2). The chemical shifts of most of the various trimethylamino amino acid derivatives were dispersed enough compared to the linewidth to be identified from each other on the basis of the differences in their side-chain structure. If two resonances overlap, the sample can be spiked with the standard trimethylated amino acid that is tentatively identified as an amino-terminus and the ^{14}N NMR spectrum can be observed as to whether its resonance corresponds exactly with that of the standard. Moreover, the wider ^{14}N chemical shift dispersion of amino acid trimethylamino derivatives compared to the ^{13}C analogues (Chapter 2) may result from the greater sensitivity of the nitrogen nucleus to environmental effects. The nitrogen atom is also one atom closer to the side-chain of the amino acid residue.

Linewidths of the quaternary trimethylamino group resonances were shown to be very sensitive to changes in electronic symmetry about the quadrupolar ^{14}N nucleus (e.g., Tables 4-1 and 4-2). Table 4-1 shows the variation in linewidth of each trimethylated amino acid standard compound. The variability in linewidth of the various amino acid trimethylamino derivatives suggests that trimethylated amino groups of amino acids derived from protein can be identified on the basis of both chemical shift and linewidth measurements.

The resonance of tetramethylammonium chloride was observed to be the sharpest, which resulted from a dipolar relaxation mechanism as opposed to a quadrupolar relaxation mechanism. The near perfect symmetry surrounding the nitrogen atom of tetramethylammonium ion of the chloride salt, its relative small size, and nondissociating groups at high pD 11.4 are factors that contribute to narrow linewidths (see theory). Sharp resonances were observed for trimethylamino groups of Ala, Gly, and Lys, but were all broader than that observed for tetramethylammonium chloride due to the distortion of the perfect electronic symmetry by the adjacent alkyl moiety in each case. The presence of bulkier side-chains in trimethylamino Ile and Val appears to lead to a more effective disruption of electronic symmetry. Even broader lines were observed for amino acid trimethylamino derivatives containing more complicated side-chains (e.g., trimethylamino Arg, Glu, His, Met, Phe, Ser, Thr, Trp, and Tyr). The largest linewidths corresponded to those for the two resonances of methylated cystine, which presumably contains both N-trimethylcystine and N,N'-hexamethyl-cystine. This interpretation is in keeping with the findings in the literature (see Introduction) that quaternized amino groups give rise to sharp resonances at high pD values and that other

derivatives, such as mono- or di- amino acids, or underivatized amino acids are not observed in the ^{14}N NMR spectrum due to line-broadening as a result of unsymmetrical electronic distribution surrounding the nitrogen atom. Moreover, the tendency of nitrogen nucleophiles to quaternize by reaction with iodomethane (March, 1985) lends support to the interpretation that the resonances arise from trimethylammonium groups.

In the case of cystine, conformational exchange may be a factor that gives rise to broad lines (Harris, 1983c). If rotation of two cysteine units about the cystine disulphide bond is moderately fast on the NMR time scale, the observed peak will be a composite of at least two semi-resolved conformations. This would place the exchange phenomena in the near-fast exchange region (Drago, 1992f). The latter could be verified by varying (increasing) the acquisition temperature, which would increase the rate of exchange without changing the separation (in Hz) of the different conformations, and would result in sharper resonances. The magnetic field strength could be increased to increase the separation between the resonances of the different conformations, but the exchange rate would not be affected. The exchange process would become even less averaged, and the resonance would be expected to broaden (Breitmaier and Voelter, 1987c).

Slight variations in the surrounding symmetrical electronic environment of the ^{14}N nucleus in some of the trimethylamino groups of amino acids derived from protein hydrolysates resulted in significant line broadening, and can be used as an aid to make peak assignments. An example of the latter was the very broad peak observed for trimethylamino phenylalanine derived from methylated insulin hydrolysate compared to the corresponding standard (Figure 4-5b). This increase in linewidth can be explained from the ^{13}C NMR spectroscopic results of ^{13}C-methylated insulin hydrolysates

(Chapter 2). Incomplete hydrolysis of insulin was observed. Most peptide bonds hydrolyze at similar rates, but the peptide bonds between the large nonpolar amino acid residues, in particular Val, Leu and Ile, require longer hydrolysis times since hydrolysis is hindered sterically by the bulky side-chains (Creighton, 1993b). Moreover, the trimethylamino group of the N-terminal Phe residue provides additional steric hindrance, which resulted in incomplete hydrolysis of ^{13}C-methylated insulin. The trimethylamino Phe-Val dipeptide was observed to be present in greater proportion than the trimethylamino Phe, even after an eight-day hydrolysis time. The line broadening is likely due to the disruption in electronic symmetry around the nitrogen nucleus as a result of the bulky side-chain of Val.

For the sake of argument, the linewidths for N^{α}-trimethylamino glycine, N^{ϵ}-trimethylamino lysine and N^{α}-trimethylamino phenylalanine standards were 1.3 Hz, 1.8 Hz, and 8.5 Hz, respectively. Clearly, the electronic environment surrounding the ^{14}N nucleus of N^{α}-trimethylamino glycine is relatively the most symmetrical of the three amino acid trimethylamino derivatives. The electronic structure surrounding the ^{14}N nucleus of N^{ϵ}-trimethylamino lysine is slightly less symmetrical than that for N^{α}-trimethylamino glycine possibly due to side-chain bending. Moreover, the aromatic ring of N^{α}-trimethylamino phenylalanine may provide some distortion of the electronic distribution surrounding the ^{14}N nucleus via ring current anisotropy. This would explain the slightly higher linewidth at half-height for this trimethylated amino acid derivative standard. Furthermore, it is interesting to note that the chemical shift of the N^{α}-trimethylamino phenylalanine is more downfield than that of the N^{α}-trimethylamino glycine or the N^{ϵ}-trimethylamino lysine. Ring current anisotropy may be one of the

factors that contributes to the difference in chemical shift from tetramethylammonium chloride between the three trimethylated amino acid derivatives.

The results of this study lend convincing support to the use of ^{14}N NMR spectroscopy as a tool to identify and characterize trimethylamino derivatives of protein. The methodology presented herein is easily applicable, cost efficient, and allows for the retention of the native protein structure if required. Table 4-1 clearly illustrates that trimethylammonium derivatives of amino acids are sensitive to microenvironment. The work of Zhang *et al.* (1994) has shown that such derivatives are also sensitive to molecular motions. Hence, linewidth measurements and dynamic NMR experiments can form the basis of an approach to study protein structure-function relationships. Further work is necessary to study derivatized and unhydrolyzed protein and the appropriate trimethylamino peptide standards.

The major drawback of the ^{14}N NMR spectroscopic approach for the use in the area of protein characterization is that it is significantly less sensitive as compared to derivatized protein labeled with 100% ^{13}C-iodomethane, as was clearly illustrated by significantly longer acquisition times for unenriched samples (e.g., tens of thousands of scans for ^{14}N NMR samples vs 600 scans for ^{13}C-NMR samples, which would take more than a day to acquire as opposed to 40 min, respectively). Thus, if structure-function studies are to be carried out, obviously the protein is desired to remain intact, and solubility constraints may limit the technique. However, as demonstrated in Chapter 3, the extent of methylation can be controlled, which can limit the extent of insolubility observed, and thus facilitate structure-function studies. Moreover, the sensitivity

problem means that large amounts of protein are required, which simply may not be readily available.

5. Acknowledgement

This research was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

6. References

Akitt, J. W. (1972) In *Annual Reports on NMR Spectroscopy*, (Mooney, E. F., ed.), Academic Press, **5A:** 466-467.

Bailey, L. (1964) *Techniques in Protein Chemistry*, Elsevier Publishing Company, p. 99.

Berglund, H., Baumann, H., Knapp, S., Landenstein, R. and Hard, T. (1995) J. Am. Chem. Soc., **117:** 12883-12884 and references therein.

Blomberg, F., Maurer, W., Ruterjans, H. (1976) *Proc. Natl. Acad. Sci. USA*, **73(5):** 1409-1413.

Breitmaier, E., and Voelter, W. (1987). In *Carbon-13 NMR Spectroscopy: High-Resolution Methods and Applications in Organic Chemistry and Biochemistry*, 3rd ed., VCH Publishers, New York. (a) p. 177; (b) p. 164; (c) p. 128-129.

Brevard, C. and Granger, P. (1981) *Handbook of High Resolution Multinuclear NMR*, John Wiley and Sons, New York, p. 104.

Cohen, E.A., Shiller, A.M., Chan, S.I., Manatt, S.L. (1975) *Organic Magnetic Resonance*, **7:** 605-609.

Cooper, R.A., Lechter, R.L., and Roberts, J.D. (1973) *J. Am. Chem. Soc.*, **95:** 3724.

Creighton, T. E. (1993). *Proteins: Structure and Molecular Properties*, 2nd ed., W.H. Freeman and Company. (a) p. 6; (b) p. 29.

Darbre, A. (1986). *Practical Protein Chemistry* (Darbre, A., ed.), Wiley, New York, p.236.

Dayhoff, M. O., ed. (1972) Atlas of Protein Sequence and Structure, Vol. 5, The National Biomedical Research Foundation, Silver Springs, Maryland.

Delange, R. J., Glazer, A. N., and Smith, E. (1969) *J. Biol. Chem.* **244:** 1385-1388.

Drago, R. S. (1992) Physical Methods for Chemists, 2nd edn., Saunders College Publishing, New York. (a) p. 319; (b) p. 326-328; (c) p. 257-258; p. 290; (d) p. 211; (e) 241; (f) 290-295.

Drakenberg, T. and Forsen, S. (1983) in *The Multinuclear Approach to NMR Spectroscopy* (Lambert, J.B. and Riddell, F.G., Eds.), Nato Advanced Study Institute, D. Reidel Publishing Co., Stirling, pp.405-444.

Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd ed., W.H. Freeman and Company, New York, pp. 33-36.

Fried, B. and Sherma, J. (1994) in *Thin-Layer Chromatography: Techniques and Applications*, 3rd Edition, Chromatographic Science Series, Vol. **66**, Ch.16.

Galanter, W.L. and Labotka, R.J. (1991) *Biochimica et Biophysica Acta*, **1079:** 146-151.

Gerothanassis, I.P., Karayannis, T., Sakarellos-Daitsiotis, M., Sakarellos, C., Marraud, M. (1987) *J. Magnetic Resonance*, **75:** 513-516.

Greenstein, J. P. and Winitz, M. (1961) In *Chemistry of the Amino Acids*, John Wiley and Sons, Inc., U.S.A., Vol. **3**, pp. 2764-2767.

Harris, R. K. (1983) *Nuclear Magnetic Resonance Spectroscopy: A Physicochemical View*, Pitman Publishing Inc., (a) pp. 130-137; (b) pp. 67-68; pp. 85-93; (c) 118-130.

Hirs, C. H. W. (1956) *J. Biol. Chem.*, **219:** 611.

Jenks, G.J. (1971) *J. Chem. Phys.*, **54:** 1558.

Jork, Funk, Fischer, and Wimmer (1990) in *Thin-Layer Chromatography: Reagents and Detection Methods*, Vol. **1b**, pp. 290-295, pp. 358-364.

Kaplan, H., Hefford, M.A., Chan, A. M.-L., Oda, G. (1984) *Biochem. J.*, **214:** 135-143.

Karayannis, T., Gerothanassis, I.P., Sakarellos-Daitsiotis, M., Sakarellos, C., Marraud, M. (1990) *Biopolymers*, **29:** 423-439.

Knapp, D.R. (1990) in *Methods in Enzymology*, **193:** 314-329.

Mann, B.E. (1978) in *NMR and the Periodic Table*, (Harris, R.K. and Mann, B.E., eds.), Academic Press Inc. (London) Ltd. (a) p. 96; (b) p. 98.

March, J. (1985) *Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, 3rd Ed., John Wiley & Sons, p. 364.

McLean, W. F. H. and Jewers, K. (1972) *J. Chromatogr.*, **74:** 297-302.

Mooney, E.F. and Winson, P.H. (1969) in *Annual Reports on NMR Spectroscopy*, (Mooney, E.F., Ed.), Academic Press Inc. (London) Ltd. **Vol. 2.**

Muller, H. and Eckert, H. (1989) *J. Chromatogr.*, **479:** 452-458.

Proctor, W.G., Yu, F.C. (1950) *Physical Reviews*, 77: 716-717.

Reisse, J. (1982) In *The Multinuclear Approach to NMR Spectroscopy* (Lambert, J. B. and Riddell, F. G., Eds.), D. Reidel Publishing Company, Ch. 4.

Reuvers, Th. B. A., Ortiz, G., Ramos, M., and De Pozuelo, M. (1989) *J. Chromatogr.*, **467:** 321-326.

Richards, R.E. and Thomas, N.A. (1975) *J. Chem. Soc., Perkin Trans.*, **2:** 368.

Sheinblatt, M. and Gutowsky, H.S. (1976) *J. Am. Chem. Soc.*, **86:** 4814.

Snyder, L. R., Palamareva, M. D., Kurtev, B. J., Viteva, L. Z., and Stefanovsky, J. N. (1986) *J. Chromatogr.*, **354:** 107-108.

Taralp, A. and Kaplan, H. (1997) *Journal of Protein Chemistry*, **16:** 183-193.

Tyihak, E., Ferenczi, S., Hazai, I., Zoltan, S., and Patthy, A. (1974) *J. Chromatogr.*, **102:** 257-262.

Tzalmona, A. and Loewenthal, E. (1974) *The Journal of Chemical Physics*, **61(7):** 2637-2639.

Vath, J.E., Zollinger, M., and Biemann, K. (1988) *Fresenius Z. Anal. Chem.*, **331:** 248-252.

Witanowski, M. and Webb, G.A. (1972) in *Annual Reports on NMR Spectroscopy*, (Mooney, E.F., Ed.), Academic Press Inc. (London) Ltd. **Vol. 5A.** (a) p. 396-398; (b) p. 409-410; (c) p. 426; (d) 435.

Witanowski, M., Stefaniak, L. and Webb, G.A. (1981) In *Annual Reports on NMR Spectroscopy*, (Webb, G.A., Ed.), Academic Press Inc. (London) Ltd. **Vol. 11B,** (a) p. 118; (b) p. 234; (c) pp. 257-261.

Zaks, A. and Klibanov, A.M. (1984) *Science*, **224:** 1249-1251.

Zaks, A. and Klibanov, A.M. (1988) *J. Biol. Chem.*, **263**: 3194-3201.

Zhang, M., Huque, E. and Vogel, H.J. (1994) *J. Biol. Chem.*, **269(7)**: 5099-5105.

Appendix 4-1

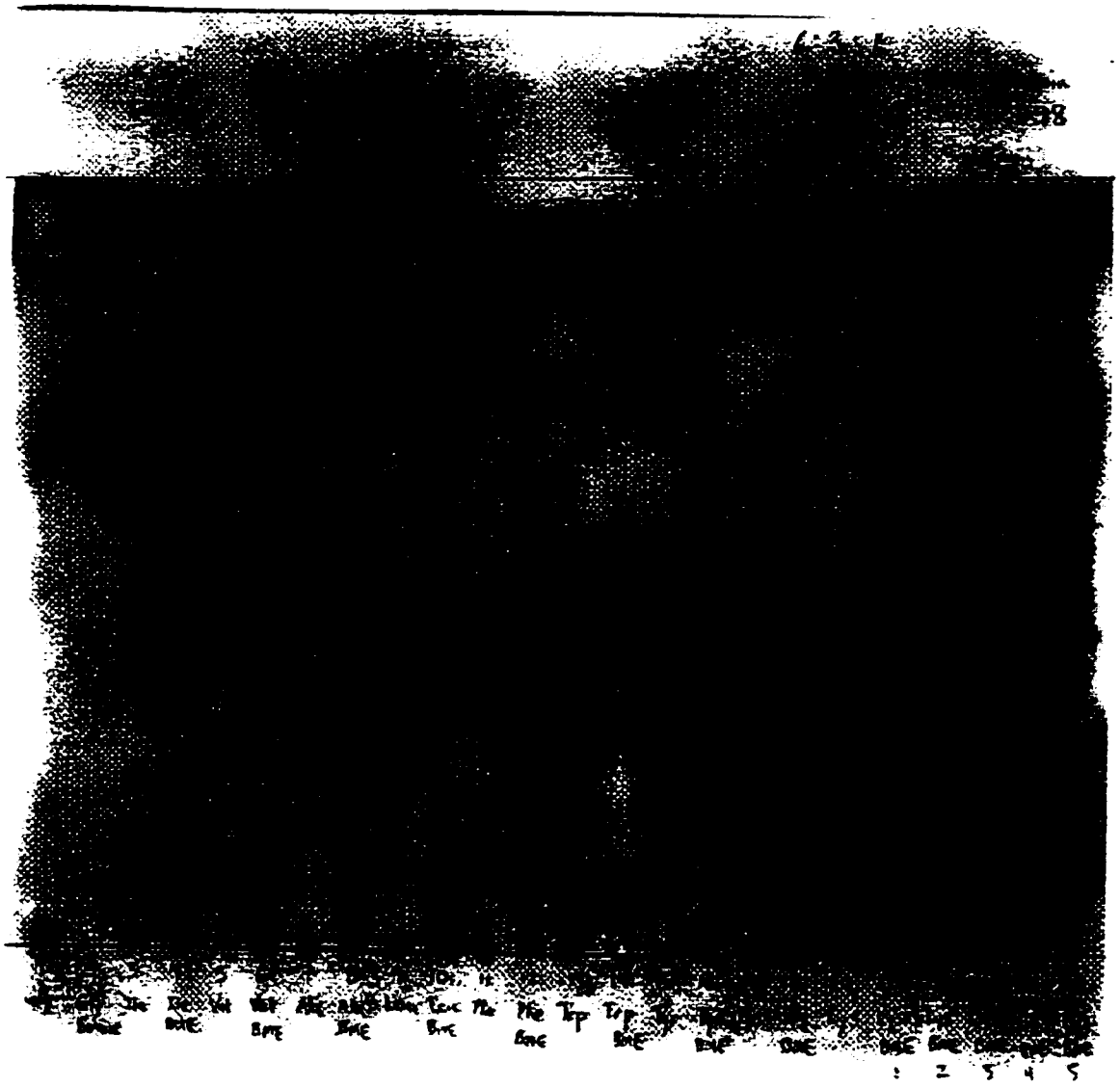


Figure 4-1(a)A Thin-layer chromatographs of quaternary ammonium derivatives of the amino acids (betaines). Sorbent: basic alumina; solvents: chloroform-methanol-ammonia (6:3:1); run time: 1.5 h.

Figure 4-1(b)A Thin-layer chromatographs of quaternary ammonium derivatives of the amino acids (betaines). Sorbent: basic alumina; solvents: chloroform-methanol-ammonia (6:3:1); run time: 1.5 h.

Chapter 5

High Specific Activity of Tracer Molecules Prepared *In Vacuo* with ^{14}C-Iodomethane

1. Introduction

Tracer studies are responsible for many advances in biochemistry (Means and Feeney, 1971; Lehninger, 1975; Walsh, 1979; Darbre, 1986; Fersht, 1985). Such studies have been among the powerful means of elucidating reaction mechanisms, biosynthetic pathways, and mechanisms of action of bioactive peptides and proteins. Tracer studies often require that bioactive peptides or proteins be labeled with high specific radioactivity, while maintaining biological activity. The latter condition often limits the incorporation to trace amounts of reagent, which places an upper limit on the specific radioactivity achievable. Radioactive ^{125}I is a commonly employed radiolabel, however it is a hazardous tracer label because it is volatile and very radioactive (Simpson *et al.*, 1979). Radioactive iodine (^{125}I) has been used for labeling of tyrosyl side-chains because of its high specific radioactivity. Unfortunately, the labeling of protein with iodine often perturbs bioactivity, in many cases nonreversibly, by disrupting native structure. The role of tyrosyl side-chains in mediating structure is known (Means and Feeney, 1971). Radioactive ^{14}C and ^{3}H are less hazardous nuclei and their incorporation into biologically relevant molecules for use in tracer studies is less disruptive to native protein structure. Taralp (1997, Doctoral Thesis) has shown that methylation of the insecticidal protein from *Bacillus thuringiensis* in aqueous medium results in the maintenance of biological activity (toxicity) up to 4 h reaction time, indicating that protein can tolerate substantial reaction with iodomethane before structure is

compromised. In this chapter the potential merit of the *in vacuo* chemical modification technique of Taralp and Kaplan (1997) is examined for the ^{14}C-methylation of protein intended for use in tracer studies in the determination of structure-function relationships. The model proteins used in this study were the insecticidal toxin isolated from *B. thuringiensis* crystals and the anabolic hormone insulin.

1.1 The Importance of Studying *Bacillus thuringiensis* Toxin

The gram-positive, spore-forming bacterium *B. thuringiensis* produces insecticidal crystalline inclusion body proteins, or δ-endotoxins, which accumulate during sporulation (Bulla *et al.*, 1980; Höfte and Whiteley, 1989). The protoxin crystal proteins vary in molecular mass from 130-140 kDa (Andrew *et al.*, 1987; Höfte and Whiteley, 1989); and in the crystal, are linked by a network of symmetrical interchain disulfide bonds (Bietlot *et al.*, 1990). Upon ingestion by susceptible insect larvae, the parasporal inclusion dissolves in the alkaline midgut environment; the crystal proteins are released by cleavage of the disulfide linkages and activated by proteolytic cleavage to 60-70 kDa toxin derived from the N-terminal portion of the molecule (Adang *et al.*, 1985; Bietlot *et al.*, 1989) by larval gut proteinases (Andrews *et al.*, 1987; Fast, 1981). The toxins then interact with the midgut apical membranes at specific receptors, disrupting membrane integrity (cell lysis) and eventually leading to insect death (Jaquet *et al.*, 1987; Hofmann *et al.*, 1988; Lüthy and Ebersold, 1984).

The precise mechanism of toxin action on the lytic mechanism is not known. The current model of *B. thuringiensis* toxin action involves ion pore formation that leads to osmotic balance disruption in the midgut epithelial layer (Gill *et al.*, 1992; Knowles and Dow, 1993; Haider and Ellar, 1987). This selective action on midgut epithelium is

facilitated by the presence of high-affinity toxin binding sites (receptors) in the brush border membranes of midgut columnar cells. The insecticidal activity of *B. thuringiensis* toxins is generally correlated with toxin binding to the brush border membrane vesicles isolated from midguts of susceptible insects, while nonsusceptible insects usually lack specific binding sites (Hofmann *et al.*, 1988; Van Rie *et al.*, 1990). The many strains of *B. thuringiensis* that have been identified differ greatly in their insecticidal activities toward various host larvae (Dulmage, 1981; Aronson *et al*, 1986; Andrews *et al.*, 1987). The differences in the toxicity of the various toxins may correspond to variances in the binding site concentrations or binding affinities to the toxin binding sites (Cowles *et al.*, 1995). Investigations are ongoing to identify and purify the toxin binding proteins, and to characterize the structure-function relationship between these receptors and the insecticidal *B. thuringiensis* toxins. Radioactively labeled toxin possessing high specific activity is required to measure the interaction of the latter with intact target cells or with isolated intact or solubilized membrane preparations derived from such cells. In the present study, the merit of the *in vacuo* protocol to incorporate high levels of radioactivity into *B. thuringiensis* toxin was investigated.

1.2 The Importance of Studying Insulin

Insulin is a polypeptide hormone that plays a key role in cellular metabolism and development (Houslay and Heyworth, 1983; Stryer, 1988). Insulin acts by binding to receptors in the plasma membrane of target cells. The insulin receptor has tyrosine kinase activity when the hormone is bound, and catalyzes the phosphorylation of tyrosine residues in target proteins (Stryer, 1988). The mechanism of insulin action is not fully understood. Investigations are ongoing to identify and purify the insulin binding

proteins, and to characterize the structure-function relationships between these receptors and insulin. Binding studies require the use of radiolabeled insulin with high specific activity; the latter requires some technical considerations.

Insulin binds to its receptors very tightly; the dissociation constant of the hormone-receptor complex is 0.1 nM. The high affinity of the receptor for insulin is essential because the insulin level in the blood is low (0.1 nM). Moreover, the number of insulin binding sites found on any given cell is very small (e.g., approximately 10^4 insulin receptors per fat cell (Cuatrecasas, 1971); one receptor per square micrometer of plasma membrane of fat cells, which corresponds to one receptor per 10^6 phospholipid molecules (Stryer, 1988)). The small number of sites and the low concentrations over which range measurements must be made necessitate the use of highly sensitive radioisotopic techniques. Molecules possessing specific activities in the range 10 to 20 Ci/mmol (attainable with ^{3}H-labeled molecules derivatized in aqueous solution) would restrict measurements (e.g., in the fat cell, with 10^6 cells per sample and 10^4 binding sites per cell) to a maximum binding of a few hundred counts per minute above background. Since experiments with large quantities of cells or membranes are impractical, it is common to work with ^{125}I- or ^{131}I-labeled or comparable derivatives that possess specific activities in the range of 1000 to 3000 Ci/mmol. Such iodinated derivatives of insulin have proved most useful both in radioimmunoassay and in receptor studies. The merit of the *in vacuo* derivatization protocol of Taralp and Kaplan (1997) was investigated to radiolabel insulin to a high degree of specific activity using ^{14}C-iodomethane; the use of ^{3}H-iodomethane may yield the required level of specific activity for binding studies.

2. Materials and Methods

2.1 Degree of ^{14}C-Iodomethane Incorporation into *Bacillus thuringiensis* Toxin Functional Groups with Time via *In Vacuo* Derivatization

B. thuringiensis toxin was prepared via the published procedure of Bietlot *et al.* (1989). Lyophilized *B. thuringiensis* toxin (*Bt*-T_1) (100 µg; 1.5 nmol) was reacted *in vacuo* at LpH 10 according to the technique of Taralp and Kaplan (1997) at various times (0, 0.5, 1, 2, 4, 8, 24, and 48 hr) at 75°C with 20 µL of a radioactive stock reagent containing ^{14}C- and ^{12}C-iodomethane in anhydrous octane. The stock reagent contained 245 µL ^{14}C-iodomethane [(specific activity: 56 mCi/mmol; 1 mCi in 1 mL anhydrous octane; 1 mCi contains 17.9 µmol; ^{14}C-iodomethane concentration: 17.9 µmol/mL) in anhydrous octane (4.39 µmol)], 24.5 µL cold (16.07 M) ^{12}C-iodomethane (393.72 µmol). and 220.5 µL anhydrous octane (= 490 µL total volume). A 20 µL aliquot of this radioactive stock reagent contains 16.25 µmol ^{14}C/^{12}C-iodomethane [20 µL of stock reagent x 398.10 µmol/490 µL = 16.25 µmol ^{14}C/^{12}C-iodomethane added per sample tube (0.179 µmol ^{14}C-iodomethane and 16.07 µmol ^{12}C-iodomethane)]. Therefore, the reagent was added in excess to the protein, which allows for the assumption to be made that the reagent is not decreasing with time. The reaction tubes were evacuated after the methylation reactions, and each sample was taken up in 1 mL distilled water. A 100 µL aliquot of each sample was placed in a scintillation vial and 5 mL of scintillation cocktail was added. Samples were counted for 1 hr using a ^{14}C/^{3}H-wide open window. Methylated *Bt*-T_1 samples were dialyzed twice against 4 L distilled water (MWCO 3500) and lyophilized in preparation for cell bioassays.

2.2 Trace-Labeling of *Bacillus thuringiensis* Toxin with ^{14}C-Iodomethane: Degree of ^{14}C-Incorporation via *In Vacuo* Derivatization

Lyophilized *Bt*-T_1 (0.5 mg; 7.5 nmol) was lyophilized in a hydrolysis tube out of 10 mM sodium metaborate buffer (1 ml), at pH 7.5 or 10. The tube was narrowed by flame and a solution of ^{14}C-iodomethane (0, 5, 10, 20, or 50 µl [56 mCi/mmole: 1 mCi in 1 ml octane]) was added. The vessel was sealed *in vacuo* and incubated at 75°C for 24 h. Heat and no heat control samples were also investigated. A 15 µg aliquot (30 µL aliquot out of 1000 µL sample in distilled water) was used for analysis by scintillation counting (1 h using a ^{14}C/^{3}H-wide open window) and a 25 µg aliquot (100 µL/1000 µL in distilled water were lyophilized and taken up in 50 µL distilled water, half of which was added to 20 µL SDS-sample buffer (Bollag and Edelstein, 1991) and 5 µL 0.2 M CAPS (3-cyclohexylamino-1-propanesulphonic acid) buffer, 1% β-mercaptoethanol, pH 10.5) was used for analysis by SDS-PAGE. Samples for LpH 7.5 and LpH 10 were run on a BIO-RAD slab gel apparatus using separate 12% homogeneous gels. Each gel was then sliced at the location of the native *Bt*-T_1 marker lanes and the portion of the gel containing the radiolabeled *Bt*-T_1 samples was allowed to air dry, and was subsequently subjected to autoradiography for 1 wk. Methylated *Bt*-T_1 samples were dialyzed twice against 4 L distilled water (MWCO 3500) and lyophilized in preparation for cell bioassays.

2.3 Cell Bioassays of Methylated *Bacillus thuringiensis* Toxin Samples

Quantification and CF-1 insect cell bioassays of ^{14}C-methylated and control toxin were performed according to published procedures (Choma and Kaplan, 1990) by Dr. Ross E. Milne of the Forest Management Institute, Forestry Canada. Lyophilized samples were incubated for 18 h in 0.1 M CAPS buffer (1 mL), pH 10.5 and centrifuged

(10K, 10 min). An aliquot of the supernatant was quantified at 595 nm using the dye binding method of Bradford (1976) and the remainder was tested as a series of dilutions to examine the threshold response.

2.4 Trace-Labeling of Insulin with ^{14}C-Iodomethane: Degree of ^{14}C-Incorporation into Insulin Functional Groups via *In Vacuo* Derivatization

Insulin (1 mg) in 10 mM sodium metaborate buffer (1 ml), at pH 9 was lyophilized in a hydrolysis tube. The tube was narrowed by flame and a solution of ^{14}C-iodomethane [0, 10, 20, 75 or 150 µl (56 mCi/mmole: 1 mCi in 1 ml octane)] was added. The vessel was sealed *in vacuo* and incubated at 75°C for 12 h. Heat and no heat control samples were also investigated. Samples (8.3 µg) were analyzed by scintillation counting [50 µL/600 µL x 100 µg (in 100 µL aliquot of 1000 µL sample) = 8.3 µg] for 1 h using a ^{14}C/^{3}H-wide open window. Two 2 µg fractions per sample [4 µL aliquot loaded onto gel from a 20 µL stock solution of derivatized insulin containing 10 µg sample (10 µL of 1 mg/mL insulin in distilled water) and 10 µL non-reducing SDS-sample buffer (Bollag and Edelstein, 1991)] were sacrificed for analyses by SDS-PAGE; two gels were run, one was stained and the other was not stained. Samples were run on a Pharmacia PhastGel™ system using 8-25% gradient gels, allowed to air dry, and the unstained gel was subjected to autoradiography for 3 wks. Methylated insulin samples were dialyzed twice against 4 L distilled water (MWCO 3500) and lyophilized in preparation for cell bioassays.

3. Results

The insecticidal protein (toxin) produced from *B. thuringiensis* was reacted *in vacuo* at LpH 10 with excess ^{14}C-iodomethane for various times. The degree of ^{14}C-incorporation into *B. thuringiensis* toxin via *in vacuo* methylation increased with time

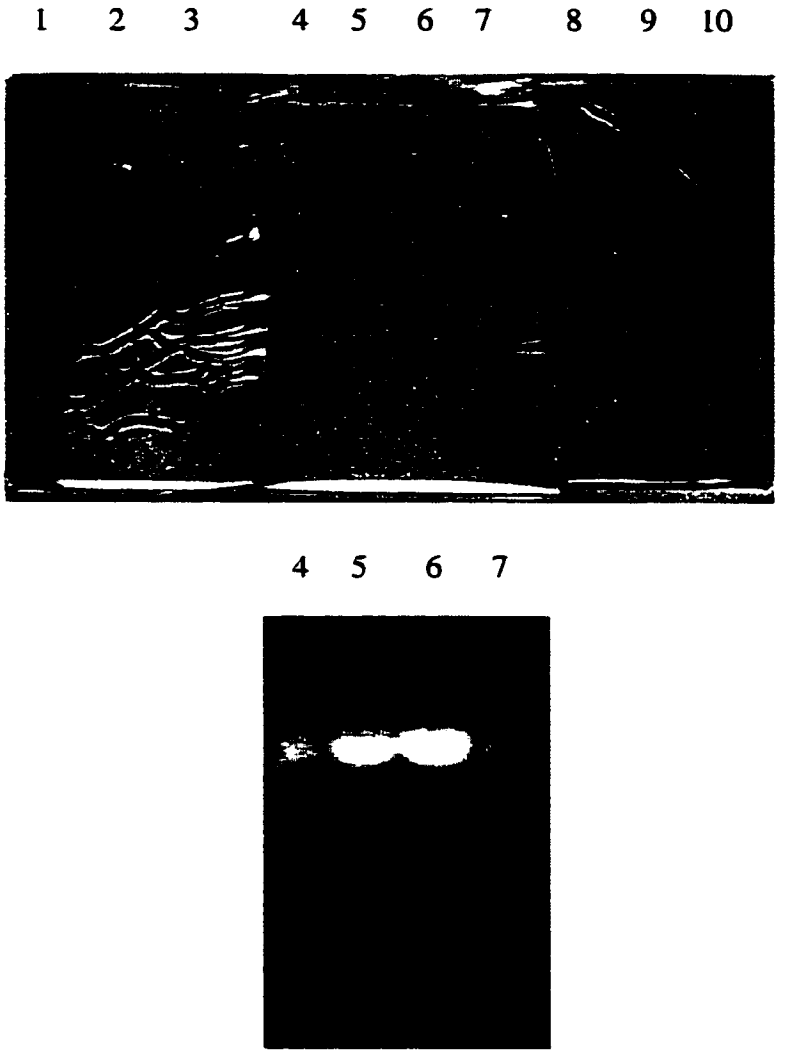


Figure 5-1. Coomassie stained (top) and autoradiographed (bottom) SDS-PAGE profile of toxin from *Bacillus thuringiensis* (*Bt*) following *in vacuo* methylation with ^{14}C-iodomethane at 75°C.

Lanes 1, 9, 10: molecular weight markers; lane 2, 3, 8: non-radiolabeled *Bt* toxin controls (0 μCi ^{14}C-CH_3I); lane 4: (5 μCi); 5, (10 μCi); lane 6: (20 μCi); lane 7: (50 μCi).

(Table 5-1). Bioactivity assays of the methylated *B. thuringiensis* toxin samples showed that activity was retained up to 8 h reaction time. These results were in contrast to those for the trace-labeled ^{14}C-methylated *B. thuringiensis* toxin samples (see below), which showed that activity was retained for samples reacted at 75°C for 24 h. The discrepancy may lie in the fact that 100 μg *B. thuringiensis* toxin were reacted in the degree of ^{14}C-incorporation experiment and 0.5 mg *B. thuringiensis* toxin were reacted in the trace-labeling experiment.

Trace labeling of *B. thuringiensis* toxin and insulin with high specific activity ^{14}C-iodomethane was carried out *in vacuo* (Figures 5-1 and 5-2, respectively). The amount of radioactivity incorporated (Tables 5-2 and 5-3, respectively) increased in direct proportion to the amount of reagent present. Losses of radiolabel were incurred after samples were dialyzed because small amounts of sample were being handled. To prevent the latter in future studies, it is reasonable to add a carrier protein such as bovine serum albumin to each sample and the losses to the radiolabeled protein would be minimized; however, a band for the carrier protein would be evident on SDS-PAGE, which may be undesirable. A micro-dialysis procedure has since been developed in our laboratory, which is another alternative to the aforementioned problem, as would be running the sample over a molecular sieve column to remove small radiolabeled contaminants, which would add to the counts if not removed.

Bioactivity assays of the trace-labeled ^{14}C-methylated *B. thuringiensis* toxin samples [with increasing amounts of ^{14}C-iodomethane (5, 10, 20, and 50 μL or μCi) for LpH 7.5 and 10 samples, 75°C, 24 h] showed that activity was retained but slightly lower levels were observed compared to the no heat and heat controls. No detrimental effect on

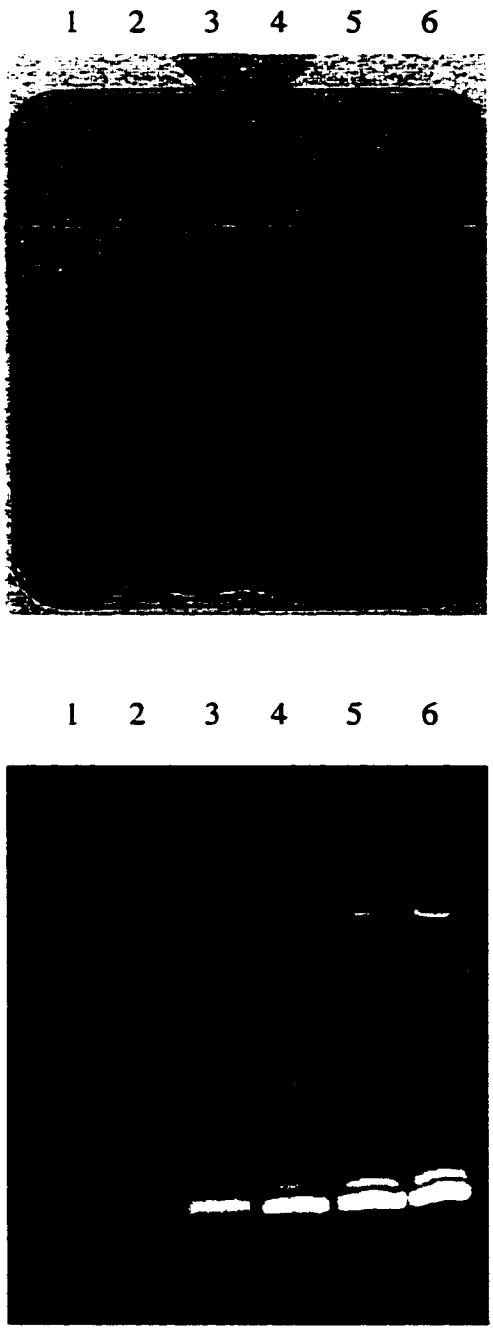


Figure 5-2. Coomassie stained (top) and autoradiographed (bottom) SDS-PAGE profile of insulin following *in vacuo* methylation with ^{14}C-iodomethane at 75°C.

Lane 1: native insulin control (0 μCi ^{14}C-CH_3I); lane 2: non-radiolabeled insulin control (0 μCi); lane 3: (10 μCi); lane 4: (20 μCi); lane 5: (75 μCi); lane 6: (150 μCi).

toxicity was observed for non-heated and heated control samples treated with anhydrous octane in place of reagent. Two samples (*B. thuringiensis* toxin, LpH 7.5 and 10, reacted with 20 μL or μCi ^{14}C-iodomethane, specific activity 56 mCi/mmol; 1 mCi/mL anhydrous octane) were further analyzed in threshold assays, whereby a dilution series was conducted and activity and counts were measured. The optimum threshold activity of *B. thuringiensis* HD-73 toxin is 0.75 ng; the no heat and heat samples were one- to two-fold dilution different at 3.125 and 1.56 ng, respectively, and were very near the optimum threshold. The *B. thuringiensis* toxin, LpH 7.5 and LpH 10 ^{14}C-methylated samples required relatively more sample (195 ng) to reach a comparable threshold level of cell bioactivity. Table 5-4 shows some published levels of radiolabel incorporation of toxin and serves as a guide to the levels required for binding assays. The low level of ^{14}C-incorporation (dpm/ng; Table 5-4) observed in our preliminary results is not surprising, since it was likely that losses were incurred during dialysis (e.g., hydrolysis of carboxyl ^{14}C-methyl esters; at the time that this research had been conducted, the observation that carboxyl groups react with iodomethane was not known). It is therefore not unreasonable to attempt further studies in order to try to improve these results.

In order to conduct a successful bioassay with *B. thuringiensis* toxin, it is required that the toxin has a specific activity of 4000 dpm/ng (4000 dpm/ng toxin x 66 x 10^3 ng toxin/nmol = 2.64 x 10^8 dpm/nmol toxin) (Lee, Young and Dean, 1995). If 20 methyl groups are incorporated per molecule of toxin using ^{14}C-iodomethane (56 nCi/nmol), the toxin will have a specific activity of 2.46 x 10^6 dpm/nmol [20 x 56 nCi/nmol toxin = 1.12 x 10^3 nCi/nmol toxin = 1.12 x 10^3 nCi x 2.2 x 10^3 dpm/nCi)/nmol toxin = 2.46 x 10^6 dpm/nmol toxin]. This corresponds to 40 dpm/ng toxin, which is not enough

for a successful bioassay (i.e., 4000 dpm/ng toxin required). However, if ^{3}H-iodomethane (25 Ci/mmol) is used, and 20 methyl groups are incorporated per toxin, the specific activity of the toxin would be 1.1 x 10^{9} dpm/nmol toxin [20 x 25 Ci/mmol toxin = 500 Ci/mmol x 2.22 x 10^{12} dpm/Ci = 1.1 x 10^{15} dpm/mmol = 1.1 x10^{12} dpm/μmol toxin = 1.1 x 10^{9} dpm/nmol toxin]. This corresponds to 16000 dpm/ng toxin [4000 dpm/ng x (1.1 x 10^{9} dpm/nmol)/(2.64 x 10^{8} dpm/nmol) ≈ 16000 dpm/ng]. Thus, assuming extensive incorporation of ^{3}H-iodomethane, the toxin would have four times the required dpm for bioassay. Thus, a less extensive chemical modification (i.e., 5 groups methylated) would be sufficient to attain the required 4000 dpm/ng toxin for bioassay using ^{3}H-iodomethane and the *in vacuo* derivatization protocol of Taralp and Kaplan (1997).

Similarly for insulin, assuming extensive modification (e.g., 10 methyl groups introduced) with ^{14}C-iodomethane (56 mCi/mmol), the incorporation would be 1.23 x 10^{6} dpm/nmol or 250 dpm/ng (0.6 Ci/mmol), which would not be enough for binding studies. The use of ^{3}H-iodomethane still comes short of the desired level of counts (1000-3000 Ci/mmol) at 5.5 x 10^{8} dpm/nmol (or 250 Ci/mmol) assuming extensive methylation. Hence, the *in vacuo* technique may prove to be more useful in derivatizing larger proteins compared to smaller ones for use in tracer or binding studies, since there are more reactable groups present in a larger protein.

Table 5-1. Degree of ^{14}C-Iodomethane[a] Incorporation into *Bacillus thuringiensis* Toxin[b] Functional Groups with Time via *In Vacuo* Derivatization

Sample	cpm per 10 µg	cpm/µg	cpm/nmol
Water blank	42	4	280
0 hr control + ^{14}MeI	248	25	1653
0.5 hr + ^{14}MeI	1024	102	6827
1 hr + ^{14}MeI	1392	139	9280
2 hr + ^{14}MeI	2281	228	15207
2 hr + ^{14}MeI	2086	209	13907
4 hr + ^{14}MeI	2584	258	17227
8 hr + ^{14}MeI	4399	434	29327
24 hr + ^{14}MeI	7301	730	48673
24 hr + ^{14}MeI	6270	627	41800
48 hr + ^{14}MeI	10448	1045	69653
48 hr + ^{14}MeI	9774	977	65160

[a] ^{14}MeI per reaction: 20 µL of stock reagent [245 µL ^{14}MeI (S.A. 56 mCi/mmol; 1 mCi in 1 mL anhydrous octane; 1 mCi contains 17.9 µmol; ^{14}MeI conc. is 17.9 µmol/mL) in anhydrous octane (4.3855 µmol) + 24.5 µL cold (16.07 M) ^{12}MeI (393.72 µmol) + 220.5 µL anhydrous octane = 490 µL total volume] = 20 µL of stock reagent x 398.10 µmol/490 µL = 16.249 µmol ^{14}C/^{12}C-MeI added per sample tube (0.179 µmol ^{14}MeI and 16.07 µmol ^{12}MeI)

[b] 100 µg (1.5 nmol) *Bt*-T_1 per reaction

Table 5-2. Trace-Labeling of *Bacillus thuringiensis* Toxin with ^{14}C-Iodomethane: Degree of ^{14}C-Incorporation via *In Vacuo* Derivatization

	Before Dialysis			After Dialysis		
Volume ^{14}MeI[a] (µL or µCi)	cpm/15 µg	cpm/µg	cpm/nmol	cpm/18.75 µg	cpm/µg	cpm/nmol
***Bt*-T_1[b] LpH 7.5**						
0						
5	8900	600	35600	4000	213	12754
10	17700	1180	70800	8700	464	27784
20	30000	2000	120000*	13830	738	44192
50	8000	533	32000	3400	181	10838
***Bt*-T_1[b] LpH 10**						
0						
5	14824	988	59296	468	25	1497
10	18176	1212	72704	673	36	2156
20	43616	2908	174464*	625	33	1976
50	6313	421	25252	150	8	479

[a] ^{14}MeI per reaction. 0-50 µL of ^{14}MeI (S.A. 56 mCi/mmol; 1 mCi in 1 mL anhydrous octane; 5, 10, 20, 50 µL of ^{14}MeI = 5, 10, 20, 50 µCi of ^{14}MeI, respecectively; 1 mCi contains 17.9 µmol; ^{14}MeI conc. is 17.9 µmol/mL)

[b] 0.5 mg/ reaction (500 µg; 7.5 nmol)

* Assuming a counting efficiency of 90%, 1 nCi would produce 2 x 10^3 count per minute (cpm). At 56 nCi/nmol specific radioactivity, 1 nmol of incorporated methyl groups from iodomethane should produce 1.12 x 10^5 cpm. It follows that if 1 nmol of *Bt*-T_1, LpH 7.5 generated 1.2 x 10^5 cpm, approximately 1 methyl group is incorporated on average, into every molecule of *Bt*-T_1 (112000 cpm/120000 = 0.93). If 1 nmol of *Bt*-T_1, LpH 10 generated 1.74 x 10^5 cpm, approximately 1.6 (or 2) methyl groups are incorporated on average, into every molecule of *Bt*-T_1 (112000 cpm/174464 = 0.64).

Table 5-3. Trace-Labeling of Insulin with ^{14}C-Iodomethane: Degree of ^{14}C-Incorporation into Insulin Functional Groups via *In Vacuo* Derivatization

Insulin[b] LpH 9	Before Dialysis			After Dialysis		
Volume ^{14}MeI[a] (µL or µCi)	cpm/30 µg	cpm/µg	cpm/nmol	cpm/8.3 µg	cpm/µg	cpm/nmol
blank	40	1.3	6.5	33	4	20
0	36	1.2	6	33	4	20
10	20000	667	3335	1740	209	1045
25	48000	1600	8000	3500	420	2100
75	68000	2267	11335	5500	660	3300
150	92000	3067	15335*	7200	864	4320

[a] ^{14}MeI per reaction: 0-150 µL of ^{14}MeI (S.A. 56 mCi/mmol; 1 mCi in 1 mL anhydrous octane; 0-150 µL of ^{14}MeI = 0-150 µCi of ^{14}MeI, respectively; 1 mCi contains 17.9 µmol; ^{14}MeI conc. is 17.9 µmol/mL)
[b] 1 mg/ reaction (1000 µg; 200 nmol)
* Assuming a counting efficiency of 90%, 1 nCi would produce 2 x 10^3 count per minute (cpm). At 56 nCi/nmol specific radioactivity, 1 nmol of incorporated methyl groups from iodomethane should produce 1.12 x 10^5 cpm. It follows that if 1 nmol of insulin generated 1.53 x 10^4 cpm, only 1 methyl group is incorporated on average, into 1 out of every 7 molecules of insulin (112000 cpm/15335 = 7.3).

Table 5-4. Published Levels of Radiolabel Incorporation of Toxin – A Guide to Levels Required for Binding Assays

Research Group	dpm/µg	dpm/ng
Lu and Adang 762 Ci/mmol input toxin ^{125}I	2.5 x 10^6	2500
Cowles, Yunovitz, Charles, and Gill 170-200 Ci/mmol ^{125}I	0.55-0.65 x 10^6	550-650
Lee, Young and Dean 1.5-2 µCi/µg ^{125}I	4.4 x 10^6	4400
English 2000-4000 dpm/ µg toxin ^{125}I		4
^{3}H-leu fermentor	8600	8.6
Vakos/Kaplan	40	0.04-0.3 (after dialysis)

Require at least 4000 dpm/ng for cell bioassays (Lee, Young and Dean). 1 Ci = 2.22 x 10^{12} dpm

4. Discussion

Bioactive peptides or proteins are typically radiolabeled for use in tracer studies for the purpose of elucidating structure-function relationships. A high degree of incorporated specific radioactivity is desired for tracking these biomolecules in tracer studies, but the upper limit of incorporation is that which results in the maintenance of bioactivity. The use of ^{14}C and ^{3}H nuclei as opposed to ^{125}I restricts the level of specific radioactivity attainable, however the benefit of maintaining structural integrity (see Introduction) makes these nuclei appealing and worth studying. Also, these nuclei have longer half-lives, which allows for the use of labeled material over a longer period of time.

Less disruptive procedures often focus on the labeling of amino groups instead of tyrosyl side-chains, and it followed, therefore, that ^{14}C-iodomethane should be explored as a possible alternate reagent to $Na^{125}I$. In water, the labeling would be most inefficient since most of the trace reagent is hydrolyzed. Previous investigators have attempted to counteract this fundamental problem by diluting the radiolabeled reagent with normal reagent. Unfortunately, the presence of excess normal reagent limits the maximum specific radioactivity incorporated and can lead to extensive and disruptive modification. When carried out *in vacuo*, however, dilution of ^{14}C-iodomethane was not necessary. Figures 5-1 and 5-2 and Tables 5-2 and 5-3 show that high specific radioactivity could be incorporated into *B. thuringiensis* toxin and insulin using neat ^{14}C-iodomethane and the *in vacuo* procedure. The gels (Figures 5-1 and 5-2) immediately showed that for equal amounts of *B. thuringiensis* toxin and insulin (top), the degree of incorporation of radioactivity increased with the amount of reagent delivered (bottom). More striking,

however, was the fact that autoradiography could be carried out with such success, and this suggested that the total radioactivity present was significant. The fact remained, however, that the incorporation was still a trace, and this was readily apparent from the results of scintillation counting. Assuming a counting efficiency of 90%, 1 nCi would produce 2 x 10^3 count per minute (cpm). At 56 nCi/nmole specific radioactivity, 1 nmole of incorporated methyl groups from iodomethane should produce 1.12 x 10^5 cpm. It follows that if 1 nmol of *B. thuringiensis* toxin, LpH 7.5 generated 1.2 x 10^5 cpm (Table 5-2), approximately 1 methyl group is incorporated on average, into every molecule of *B. thuringiensis* toxin (112000 cpm/120000 = 0.93). If 1 nmol of *B. thuringiensis* toxin, LpH 10 generated 1.74 x 10^5 cpm, approximately 1.6 (or 2) methyl groups are incorporated on average, into every molecule of *B. thuringiensis* toxin (112000 cpm/174464 = 0.64). Similarly for insulin, if 1 nmol generated 1.53 x 10^4 cpm (Table 5-3, bottom entry), only 1 methyl group is incorporated on average, into 1 out of every 7 molecules of insulin. As there are many reactable groups in each *B. thuringiensis* toxin or insulin molecule, it may be possible to improve incorporation, without losing biological activity, or turn to the use of a more radioactive source (i.e., with higher specific activity) such as ^{3}H-iodomethane (even though it is a weak emitter of β radiation: Darbre, 1986) for tracer and binding studies (see calculations in Results). The obvious next step is to use less diluted fresh ^{14}C-iodomethane stock (e.g., take up the new radiolabeled reagent containing 1 mCi in a smaller amount of anhydrous octane; it is also possible to purchase 5 mCi of ^{14}C-reagent) and repeat the experiments. Whatever direction is chosen, the initial results foreshadow the widespread use of *in vacuo* derivatizations for trace labeling protein.

The trace-labeling experiment with *B. thuringiensis* toxin was carried out in two separate experiments because in the initial experiment the reagent was added to the sample with anhydrous octane in the amount of the difference in volume of the ^{14}C-iodomethane reagent from 50 μL total volume. The samples of this initial experiment were not completely soluble after derivatization and were separated into supernatant and pellet. Solubility measurements were carried out [$A_{280\,nm}$ readings; $\varepsilon = 1.61$ mL/mg (Bietlot *et al.*, 1989)]. The results of the initial trace-labeling experiment also (cf. second experiment results shown above) showed that the amount of radioactivity incorporated increased in direct proportion to the amount of reagent present and bioactivity was retained (data not shown). For *B. thuringiensis* toxin, LpH 10, the progressive increase in ^{14}C-methylation coincided with an increase in solubility. These observations suggested that the charged groups of the methylated protein (trimethyl lysine or α-trimethyl amino group, dimethyl imidazolium cation) were mainly surface exposed and not crucial for the maintenance of bioactivity. The derivatized residues (lysyl, histidyl, aspartyl, glutamyl, N- and C-terminal amino and carboxyl groups, respectively) were observed to react by ^{13}C NMR spectroscopy for ^{13}C-methylated *B. thuringiensis* toxin (see Chapters 2 and 3) and also by amino acid analysis (Taralp, 1997, Doctoral Thesis; aqueous methylation: 70% lysyl, 50% histidyl residues reacted). The progressive increase in ^{14}C-methylation with an increase in solubility implies that the overall charge of the derivatized protein is more positive compared to the native protein. The latter scenario is likely since methylation of carboxyl groups was observed to occur (see Chapters 2 and 3). The phenolic side-chains of tyrosyl residues are in principle reactive at LpH 10 where they are partially deprotonated, but these side-chains were not

observed to react with ^{13}C-iodomethane (Chapters 2 and 3), indicating that these typically internal residues are protected from reaction during the *in vacuo* methylation. Taralp (1997, Doctoral Thesis) has proposed that the observed loss of toxicity (bioactivity) of aqueous-methylated *Bt* T_1 after 4 h reaction time was due to the loss of structure as a result of the disruption of tyrosyl interactions by methylation of the phenolic hydroxyl functions or a consequence of the loss of structure for other reasons, which resulted in subsequent methylation of exposed tyrosine groups. It appears that *in vacuo* derivatization compared to aqueous derivatization preserves protein structure and bioactivity because of the lack of conformational dynamic equilibria in the former protocol (and the diminished risk for proteolysis).

5. Acknowledgement

This research was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada. The author would like to especially thank Dr. Alpay Taralp for his time, guidance and mentoring during the course of the present research. The author would like to thank Dr. Ross Milne of the Forest Pest Management Institute, Forestry Canada for carrying out bioactivity assays on derivatized *B. thuringiensis* toxin samples. The author would like to thank Mr. Barry Lorbetski for technical support.

6. References

Adang, M. J., Staver, M. J., Rocheleau, T. A., Leighton, J., Barker, R. F. and Thompson, D. V. (1985) *Gene* **36:** 289-300.

Andrews, R. E., Jr., Faust, R. M., Wabiko, H., Raymond, K. C., and Bulla, L. A., Jr. (1987) *CRC Crit. Rev. Biotechnol.*, **6:** 163-232.

Aronson, A. I., Beckman, W., and Dunn, P. (1986) *Microbiol. Rev.*, **50:** 1-24.

Bietlot, H. Vishnubhatla, I., Carey, P. R., Pozsgay, M., and Kaplan, H. (1990) *Biochem. J.*, **267:** 309-315.

Bietlot, H., Carey, P. R., Choma, C., Kaplan, H., and Pozsgay, M. (1989) *Biochem. J.*, **260:** 87-91.

Bulla, L. A., Jr., Bechtel, D. B., Kramer, K. J., Shethna, Y. I., Aronson, A. I., and Fitz-James, P. C. (1980) *Crit. Rev. Microbiol.*, **8:** 147-204.

Cowles, E. A., Yunovitz, H., Charles, J.-F.., and Gill, S. S. (1995) *Appl. Environ. Microbiol.*, **61:** 2738-2744.

Cuatrecasas, P. (1971) *Proc. Natl. Acad. Sci.*, U.S.A. **68:** 1264.

Darbre, A. (1986) *Practical Protein Chemistry* (Darbre, A. Ed.) John Wiley and Sons.

Dulmage, H. T. (1981) in Microbial Control of Pests and Plant Diseases (Burgess, H. D., ed.) pp. 193-222, Academic Press Ltd., London.

Fast, P. G. (1981) in *Microbial Control of Pests and Plant Diseases* (Burgess, H. D., ed.) pp. 223-248, Academic Press Ltd., London.

Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd ed., p. 233, W. H. Freeman and Company.

Gill, S. S., Cowles, E. A. and Pietrantonio, P. V. (1992) *Annu. Rev. Entomol.*, **37:** 615-636.

Haider, M. Z. and Ellar, D. J. (1987) *Biochem. J.*, **248:** 197-201.

Hofmann, C., Vanderbruggen, H., Höfte, H, Ban Rie, J., Jansens, S. and Van Mellaert, H. (1988) *Proc. Natl. Acad. Sci. USA.*, **85:** 7844-7848.

Höfte, H. and Whiteley, H. R. (1989) *Microbiol. Rev.*, **53:** 242-255.

Houslay, M. D. and Heyworth, C. M. (1983) *TIBS*, pp. 449-452.

Jaquet, F. Hütter, R. and Lüthy, P. (1987) *Appl. Environ. Microbiol.*, **53:** 500-504.

Knowles, B. H. and Dow, J. A. T. (1993) *Bioessays*, **15:** 469-476.

Lee, M. K., Young, B. A., and Dean, D. H. (1995) *Biochem. Biophys. Res. Commun.*, **216:** 306-312.

Lehninger, A. L. (1975) *Biochemistry: The Molecular Basis of Cell Structure and Function*, 2nd ed., p. 379, pp. 461-463, Worth Publishers.

Lüthy, P. and Ebersold, H. R. (1984) in *Pathogenesis of invertebrate microbial diseases* (Davidson, E. W., ed.) pp. 235-267, Allanheld, Osmum and Co., Totowa, NJ.

Means, G. E. and Feeney, R. E. (1971) *Chemical Modificaton of Proteins*, Holden-Day.

Schwartz, J. L., Lu, Y. J., Sohnlein, P., Brousseau, R., Laprade, R., Masson, L., Adang, M. J. (1997) *FEBS Lett.*, **412:** 270-276.

Simpson, J. S. A., Campbell, A. K., Ryall, M. E. T. and Woodhead, J. S. (1979) *Nature* **279:** 646-647.

Stryer, L. (1988) *Biochemistry* pp. 994-998, Third Edn., W. H. Freeman and Co., New York.

Taralp, A. (1997) In *Aqueous and Nonaqueous Chemical Approaches to Elucidating Structure and Function in Proteins*. Ph.D. Thesis, Department of Chemistry, University of Ottawa, Ottawa, ON.

Taralp, A. and Kaplan, H. (1997). *J. Protein Chem.* **16,** 183-193.

Van Rie, J., Jansens, S., Höfte, H., Degheele, D., and Van Mellaert, H. (1990) *Appl. Environ. Microbiol.* **56:** 1378-1385.

Chapter 6

Cholesterol Esterase:
A. Purification,
B. Characterization of Taurocholate-Induced Dimerization in Sodium Dodecylsulfate, and
C. Bile Salt-Modulated Stereoselection in the Hydrolysis of α-Tocopheryl Acetates; Effect of Ester Group Chain Length on the Rates of Hydrolysis of Synthetic α-Tocopheryl Esters

1. Introduction

Cholesterol esterase (CE) (carboxyl ester hydrolase, EC 3.1.1.13) catalyzes the hydrolysis of cholesteryl esters, vitamin E acetate (α-tocopheryl acetate), and other lipids and water insoluble substrates (Rudd and Brockman, 1984). The hydrolysis occurs in the small intestine and the enzyme is present in pancreatic juice. CE is catalytically active only in the presence of bile salts (Rudd and Brockman, 1984), from which fact arises its modern name: bile salt-activated lipase or BAL (Wang and Hartsuck, 1993). During a study of the bile salt-modulated stereoselectivity of bovine and porcine CEs (Moore *et al.*, 1994, 1995; in which this author is acknowledged for technical assistance; see Part C), it was realized that most commercial CE preparations are quite impure and, although there are many reports of purification of CE or BAL from a number of sources, including rat (Calame *et al.*, 1975; Erlanson, 1975; Hyun *et al.*, 1969; Roberts *et al.*, 1988), cow (Cox, 1990), pig (Labow *et al.*, 1983; Momsen and Brockman, 1977), human pancreas (Abouakil *et al.*, 1988; Lombardo *et al.*, 1978; Riley *et al.*, 1990) and human milk (Bläckberg and Hernell, 1981; Hernell and Olivecrona, 1974; Swan *et al.*, 1992; Wang, 1980), most of the methods are tedious, time-consuming and give low yields.

Affinity chromatography has proven to be a powerful method for isolating pure enzymes with a minimum number of steps (Cuatrecasas *et al.*, 1968). Since CEs are characterized by specific interactions with bile salts (Hyun *et al.*, 1971; Hyun *et al.*, 1972; Moore *et al.*, 1994, 1995; Vahouny *et al.*, 1965), a bile acid-derivatized support was an obvious choice for an affinity-based chromatographic purification procedure (Moore *et al.*, 1996). Indeed, a successful method for purification of human milk BAL using a bile acid affinity column eluted with cholate already has been reported (Wang, 1980). Surprisingly, there appear to be no reports of successful applications of this method to the purification of other BAL. Our attempts to purify either porcine or bovine pancreatic CEs from their crude commercial preparations using commercially available, bile acid-derivatized crosslinked agarose (Sigma) failed to yield satisfactory results. However, success has been observed in a single chromatographic step with cholate-derivatized Sepharose prepared in-house, eluting the bound esterase with a concentration gradient of buffered taurocholate (Moore *et al.*, 1996). Purification of CEs or BALs by this quick method is, in part, the focus of the present work (see Part A). During the course of purifying bovine CE, the enzyme appeared to be more heterogeneous after elution from the cholate-derivatized Sepharose affinity column with taurocholate. The unusual phenomenon observed on SDS-PAGE was characterized (Vakos *et al.*, 1997) and is, in part, the second objective of the present study.

Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a major analytical tool for establishing the purity of a protein and obtaining an estimate of its molecular mass (Hames, 1990; See and Jackowski, 1989). While the value obtained for the apparent molecular mass may be substantially in error due to anomalous

electrophoretic mobility (Lin *et al.*, 1994; Leach *et al.*, 1980; Bayreuther *et al.*, 1980), SDS-PAGE is generally regarded as a reliable measure of the homogeneity of a protein preparation. Nevertheless, the appearance of multiple bands due to the failure to dissociate polypeptide chains completely can give an impression of a heterogeneity that does not exist in the protein preparation (Bensadoun *et al.*, 1974). In such cases, it is common practice to heat the protein samples in buffer containing SDS for longer periods of time to completely dissociate the polypeptide chains prior to applying the samples to the gel (Bensadoun *et al.*, 1974; Deutsch, 1976).

The affinity-purified bovine CE of the present study yielded several higher molecular weight bands on SDS-PAGE, especially one appearing to correspond to a dimer. Instead of decreasing the amount of dimer, longer heating times actually increased the amount of dimer relative to the monomer. It appeared that heating in SDS was promoting association of monomeric subunits to form dimers rather than the expected dissociation. The second objective of the present investigation was to elucidate the factors that influence the apparent dimerization of CE on heating in SDS and to characterize the nature of the dimerization (see Part B).

The third objective of the present study deals with the bile salt-modulated diastereoselectivity of the CE-catalyzed hydrolysis of α-tocopheryl acetates (Moore *et al.*, 1995; Vakos, unpublished results; see Part C). Natural vitamin E [(*2R, 4'R, 8'R*)-α-tocopherol, or *RRR*-α-TOH], and synthetic vitamin E [(*2RS, 4'RS, 8'RS*)-α-tocopherol, or *all-rac*-α-TOH, an approximately equal mixture of eight possible stereoisomers], are available for human ingestion as esters, most usually as the acetate (see Figure 6-1) (Weiser and Vecchi, 1981). CE must hydrolyze the vitamin E ester to the free

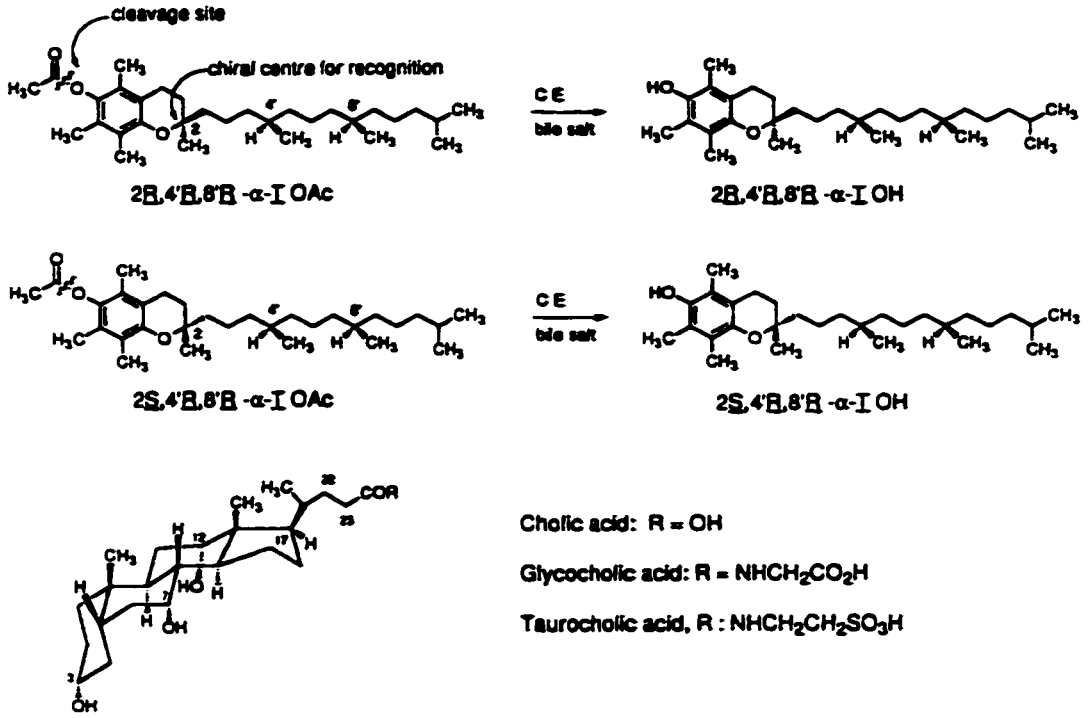


Figure 6-1. Natural vitamin E, (*2R,4'R,8'R*)-α-tocopherol (*RRR*-α-TOH), and synthetic vitamin E, (*2S,4'R,8'R*)-α-tocopherol (*SRR*-α-TOH). The natural 3α,7α,12α-trihydroxy bile salts, cholate, glycocholate, and taurocholate are cofactors or activators of CE.

α-tocopherol in order for intestinal absorption to occur. The natural 3α,7α,12α-trihydroxy bile salts, cholate, glycocholate, and taurocholate (see Figure 6-1), are particularly effective activators of CE (Lombardo and Guy, 1980; Korzenovsky *et al.*, 1960; Vahouny *et al.*, 1965; Hyun *et al.*, 1969; Erlanson, 1975; Bläckberg *et al.*, 1981; Hyun *et al.*, 1972). The obligatory role of bile salts on CE-catalyzed hydrolysis reactions is well established (Gallo-Torres, 1970a, 1970b; Mathias *et al.*, 1981a, 1981b; Muller *et al.*, 1976; Lombardo and Guy, 1979, 1980; Lombardo *et al.*, 1980a, 1980b; Korzenovsky *et al.*, 1960; Vahouny *et al.*, 1965; Hyun *et al.*, 1969; Erlanson, 1975; Rudd and Brockman, 1984; Bläckberg *et al.*, 1981). Bile salts appear to serve four functions: (a) To activate CE (Rudd and Brockman, 1984); (b) to function as detergents (Lombardo and Guy, 1980; Vahouny *et al.*, 1965; Hyun *et al.*, 1969; Erlanson, 1975; Lombardo *et al.*, 1980b; Carey and Small, 1970, 1972; Lynn *et al.*, 1982), by forming micelles (Carey and Small, 1970, 1972) and mixed micelles with dietary lipids (Carey and Small, 1970, 1972; Borgstrom, 1967; Vahouny *et al.*, 1964) together with the ester substrate which makes the ester more readily available for the CE-catalyzed reaction (Mathias *et al.*, 1981a; Mathias *et al.*, 1981b); (c) to protect CE from degradation by pancreatic proteases (Vahouny *et al.*, 1965; Hyun *et al.*, 1969; Erlanson, 1975; Hyun *et al.*, 1972); this protection apparently arises from a bile salt-induced oligomerization of the CE (Lombardo and Guy, 1980; Erlanson, 1975; Hyun *et al.*, 1972; Calame *et al.*, 1975), which produces the active form of the enzyme (Calame *et al.*, 1975); deoligomerization to reform monomer occurs if the bile salt is removed (Hyun *et al.*, 1972); (d) To protect CE from denaturation at lipid-water interfaces, probably by preventing a lipid-surface-induced conformational change which leads to deactivation (Tsujita and Okuda, 1990).

The preliminary studies of Burton, Ingold and co-workers (Zahalka *et al.*, 1991; Moore *et al.*, 1994) have uncovered a fifth role for the bile salts: an unprecedented modulation of the chiral selectivity of an enzyme. Bovine CE (BCE) was shown to catalyze the *in vitro* hydrolysis of *RRR*-α-TOAc and of an "unnatural" stereoisomer, *SRR*-α-TOAc, when the substrates were dispersed in 40 mM sodium cholate/2 mM *dl*- or *l*-dimyristoyl-phosphatidylcholine (*dl*- or *l*-DMPC) mixed micelles at 37°C in Tris buffer, pH 7.7 (Zahalka *et al.*, 1987). Noncompetitive and competitive experiments were carried out and the product α-tocopherol stereoisomers (differentiated by mass by three deuterium atoms on one of the methyl groups for the *SRR* stereoisomer) were measured by gas chromatography/mass spectrometry (GC/MS). The results showed that pure *RRR*-α-TOAc was hydrolyzed more slowly than *SRR*-α-TOAc by the BCE/sodium cholate/DMPC mixed micellar system. These *in vitro* results were inconsistent with the results of some experiments carried out using rats, which clearly demonstrated that under *in vivo* conditions the acetate of the natural stereoisomer of α-tocopherol (i.e., *RRR*-α-TOH) was hydrolyzed more extensively and, presumably, more rapidly than the acetate of the "unnatural" stereoisomer (i.e., *SRR*-α-TOH) (Ingold *et al.*, 1987). The *in vitro/in vivo* difference in chiral selectivity was later conclusively demonstrated to be due to the composition of the rat bile salts and not to any intrinsic differences in the diastereoselectivities for (α-TOAc) ester hydrolyses between CE derived from rat pancreas and CE derived from bovine pancreas (Zahalka *et al.*, 1991). Cholic acid is not present in detectable amount as the free acid in rat bile, but rather as its taurine and glycine conjugates (Norman, 1954; Subbiah *et al.*, 1969; Yousef *et al.*, 1972). Noncompetitive and competitive experiments using the sodium salts of cholic,

glycocholic, and taurocholic acids have shown that the diastereoselectivity of the BCE-catalyzed hydrolyses of α-TOAcs could be "modulated" by the bile salt used to "activate" the enzyme (Zahalka *et al.*, 1987, 1991).

In the present study, the absolute and relative rates of hydrolysis of *RRR*-α-TOAc and *SRR*-α-TOAc in noncompetitive and competitive experiments with CE in the presence of the three natural 3α,7α,12α-trihydroxy bile salts were studied (Figure 6-1). The concentrations of the bile salts and the lipid (*l*-DMPC) that interact with the bile salt to form the mixed micelles have been varied. The work regarding the bile salt diastereoselectivity of hydrolytic reactions catalyzed by CE has contributed to the more detailed study of Moore *et al.* (1995) with crude and purified (via affinity chromatography using cholate-derivatized Sepharose 4B) bovine and porcine CE. The results of the present study pertaining to the purification of CE and the characterization of the taurocholate-induced dimerization of BCE in SDS have been helpful in the overall understanding of the diastereoselectivity of hydrolytic reactions catalyzed by the CE/bile salt/lipid mixed micellar system. It was observed (Moore *et al.*, 1995) that the level of purity and the origin (i.e., cow or pig) of the CE has a much lesser influence on the degree of chiral discrimination of the enzyme (as measured by the relative rates of hydrolysis of *RRR*-α-TOAc and *SRR*-α-TOAc in competitive experiments) than does the structure of the trihydroxy bile salt used to activate the enzyme. Significantly, Moore *et al.* (1994) have shown that the origin of the enzyme is very important in the case of the only dihydroxy bile salt exhibiting a strong activating effect on a CE (Moore *et al.*, 1994). The modulating effect of the bile salt on the chiral selectivity of CE is quite

impressive in that the bond cleaved by the esterase is six bonds removed from the chiral center (see Figure 6-1).

Preliminary experiments were carried out to study the effect of the ester group chain length on the rates of hydrolysis of (synthesized) α-tocopheryl esters (*RRR*-α-**T** esters) by the CE in cholate and taurocholate bile salt/*l*-DMPC mixed micellar systems. The effect of the substrate structure on reactivity may reflect the lifetime of the acylenzyme intermediates as well as product inhibition.

PART A. Purification of Bile Salt-Activated Lipases

2A. Material and methods

Except where stated, all materials were purchased from Sigma (St. Louis, Mo.).

2A.1 Preparation of cholate-derivatized Sepharose.

Sodium cholate (500 mg), dissolved in doubly-distilled water (20 mL), was added to EAH Sepharose 4B (Pharmacia, Uppsala; 10 atom spacer arm, bead diameter 45-165 μm; 10 mL) in an Erlenmeyer flask and mixed on an orbital shaker. A solution of 1-ethyl-3-(3-dimethylamino-propyl) carbodiimide hydrochloride (500 mg) in doubly-distilled water (6 mL) was added dropwise with stirring over 5 minutes. The suspension was shaken for 24 hours. Throughout this period, 1% HCl was added periodically to maintain a neutral pH. The gel was then collected on a sintered glass filter, resuspended in phosphate buffer (100 mM, pH 8), and subsequently the buffer was filtered off. This washing procedure was repeated six times, taking care not to damage the gel particles. The gel was then washed six times in the same manner with TRIS buffer (50 mM, pH 7.5

in 80 mM NaCl containing 0.02% NaN_3). It can be stored refrigerated indefinitely in this buffer. The same buffer also is used for the affinity chromatography.

2A.2 Protein concentrations.

Protein concentrations were measured by the Lowry method (Lowry *et al.*, 1951).

2A.3 Enzyme assays.

A nonspecific assay was used to test for the presence of esterase activity in column fractions: 1.5 mL of 25 mM TRIS buffer (pH 7.5) in 40 mM NaCl was placed in a disposable UV cuvette. To this was added 50 μL of 4 mM *p*-nitrophenylbutyrate, *p*-NPB, in acetonitrile (final concentration, 1.33×10^{-4}M). An aliquot (25 μL) from a column fraction was added, and the increase in absorbance was followed at 400 nm.

Specific activity was determined using micellar cholesteryl-3-glutaric acid resorufin ester, CGRE (Beohringer Mannheim), as the substrate. Into a glass screw top vial were pipetted a solution of CGRE in dichloromethane (0.57 mM, 250 μL) and a solution of *l*-dimyristoyl phosphatidyl choline (Avanti Polar Lipids, Alabaster, Alabama) in dichloromethane (32 mM, 125 μL). The residue obtained after the solvent was evaporated under a stream of nitrogen was suspended in an aqueous solution of sodium taurocholate (0.1 M, 0.8 mL) by vortex-mixing for one minute. To this was added water (1 mL) and 0.5 M TRIS (pH 7.5)/1.5 M NaCl buffer (0.2mL). The solution was vortex-mixed for 30 seconds. The solution (1.5mL) was then placed in a disposable UV cuvette. After addition of an aliquot (*ca* 100 μL) of enzyme solution, the absorbance increase at 572 nm was followed spectrophotometrically ($\varepsilon = 60{,}000\ M^{-1}cm^{-1}$). The initial zero order rate was converted to the pseudo-first order rate by dividing by the substrate

concentration for determination of specific activity. Specific activity is reported as units/mg protein, where one unit is 1 μM CGRE hydrolyzed per second (μMs^{-1}).

2A.4 Electrophoresis.

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed with either the BioRad vertical slab gel apparatus or the Pharmacia PhastGel™ system using 10-15% gradient gels. In the former case, stock solutions for stacking and resolving gel preparations were made according to Hames (Hames, 1990) for discontinuous SDS-PAGE. All gels were stained with Coomassie blue. Samples (total volume 20 μL) were prepared in buffer that was 5 times more concentrated using the original recipe of Bollag and Edelstein (Bollag and Edelstein, 1991). Purified esterase samples were prepared in a 3:1 SDS:protein ratio to prevent the oligomerization that has been observed to occur upon sample heating (see Part B; Vakos, Kaplan and Burton, 1997).

2A.5 Affinity chromatography of porcine cholesterol esterase (PCE), bovine cholesterol esterase (BCE) and human milk bile salt-activated lipase (BAL).

Cholate-derivatized Sepharose, suspended in the TRIS storage buffer described above, was poured into 20 mL disposable plastic columns (BioRad). The gel was allowed to settle and additional gel added, if necessary, to bring the final level approximately to the 7 mL mark. The column was eluted with the same TRIS buffer (10 mL). The buffer level was allowed to drop to the top of the gel, and crude PCE (28 mg, Sigma catalogue #C-9530) in TRIS storage buffer (1 mL) was loaded onto the column. The protein was allowed to distribute through the column at a reduced flow rate of approximately 0.1 mL/minute (using a pinch clamp) over a period of approximately 10 minutes. The flow was then stopped, and the column allowed to equilibrate for

1 hour. Elution was then recommenced with the TRIS buffer at a flow rate of approximately 1 mL/minute, collecting the eluate in 2 mL fractions. The overall protein concentration was monitored by measuring the absorbance of fractions at 280 nm. This was done manually or by using a flow cell and UV detector connected to the fraction collector. Esterase activity of fractions was measured using the *p*-NPB assay. Elution with the TRIS buffer was continued until the absorbance returned to background levels (i.e. after passage of approximately 40 mL of eluent). The eluent was then changed to a buffer containing TRIS (50 mM, pH 7.5), NaCl (80 mM) and sodium taurocholate (8 mM). A slight increase in absorbance in the fractions indicated that this buffer caused elution of additional, contaminating protein. Elution was continued until the absorbance returned to background levels. A taurocholate gradient (8-40 mM) was introduced during the passage of the next 60 mL of buffer. Although a gradient mixer can be used, the same effect can be obtained using a simple mixer consisting of a siphon connecting two Erlenmeyer flasks containing the lower and higher concentration buffers, the lower concentration buffer being stirred magnetically. The esterase activity began to elute at a taurocholate concentration of approximately 25 mM, peaking and returning to baseline within the next 40 mL. Fractions that showed esterase activity were evaluated for purity by SDS-PAGE. Pure fractions were pooled and concentrated by ultra-filtration (Amicon, YM30 membrane) to a volume appropriate for storage. The purified cholesterol esterase was found to be quite stable in the final elution buffer, showing negligible loss of activity after being frozen for six month periods and undergoing freeze-thawing several times.

The same method has been successfully applied to the purification of bovine cholesterol esterase (BCE; 26 mg, Sigma catalogue # C-3766; see Figure 6-2B and

Figure 6-3(a)). Comparison of the esterase activity elution profiles of BCE purified with the in-house cholate-derivatized Sepharose 4B affinity column and the commercial Sigma cholate-derivatized agarose column revealed an improved resolution or separation and purification with the former column (cf. Figure 6-2B and Figure 6-2C).

BAL from human milk also has been purified successfully (see Figure 6-2D and Figure 6-4). The steps taken prior to performing affinity column chromatography were essentially the same as those of Wang (Wang, 1980). Cream was removed from milk by placing a total of 100 mL of milk in several 50 mL thin wall, plastic centrifuge tubes and centrifuging for 45 minutes at 17,000 rpm (37,500 g) at 4 °C in a Beckman J2-MC centrifuge using a Beckman JS-17 123 rotor. The skim milk was acidified to pH 4.7 with 6 M HCl and centrifuged for 1 h at 4 °C under conditions similar to those described already. The protein pellet was discarded and the pH of the whey adjusted to 8.5 by adding 1.0 M NaOH solution. Phenylmethylsulphonyl fluoride was added (0.4 mM final concentration) as a precaution to inhibit any proteases that may have been present. Application of a sample of the whey (9 mL) to the cholate Sepharose affinity column yielded a pure BAL, as confirmed by a single band on SDS-PAGE (apparent MW 120 kDa on a 12 % homogeneous slab gel; see Figure 6-4).

3A. Results and Discussion

The 280 nm absorbance and esterase activity elution profiles of PCE are shown in Figure 6-2A. The UV absorbance profile shows a large increase initially, as is characteristic of affinity chromatography. The components of the crude enzyme preparation that have no affinity for the cholate Sepharose come through as "wash-through", appearing at what is essentially the void volume of the column. The absence of

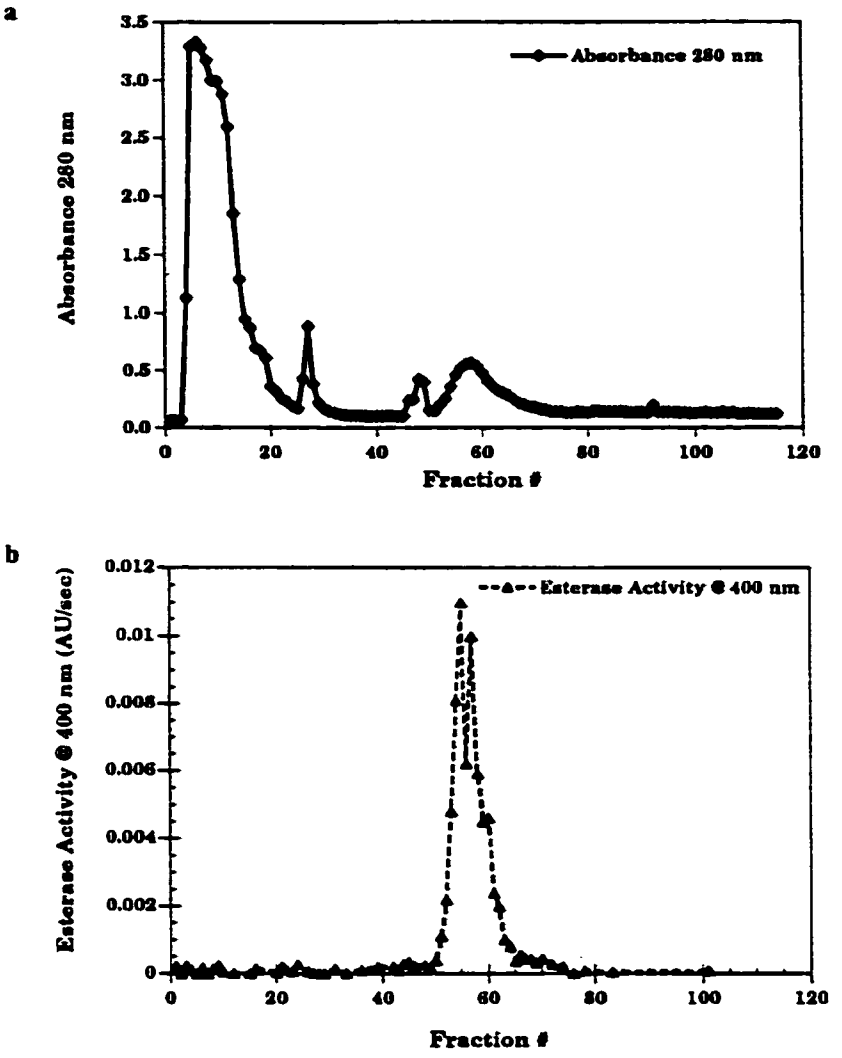


Figure 6-2A. (a) Protein elution profile as measured by absorbance at 280 nm. A switch from the initial elution buffer to 8 mM taurocholate buffer was made at fraction number (F) 21, followed by an 8-40 mM taurocholate buffer at F46. (b) Esterase activity elution profile for PCE (using in-house-prepared cholate-derivatized Sepharose 4B affinity column) as determined using *p*-NPB.

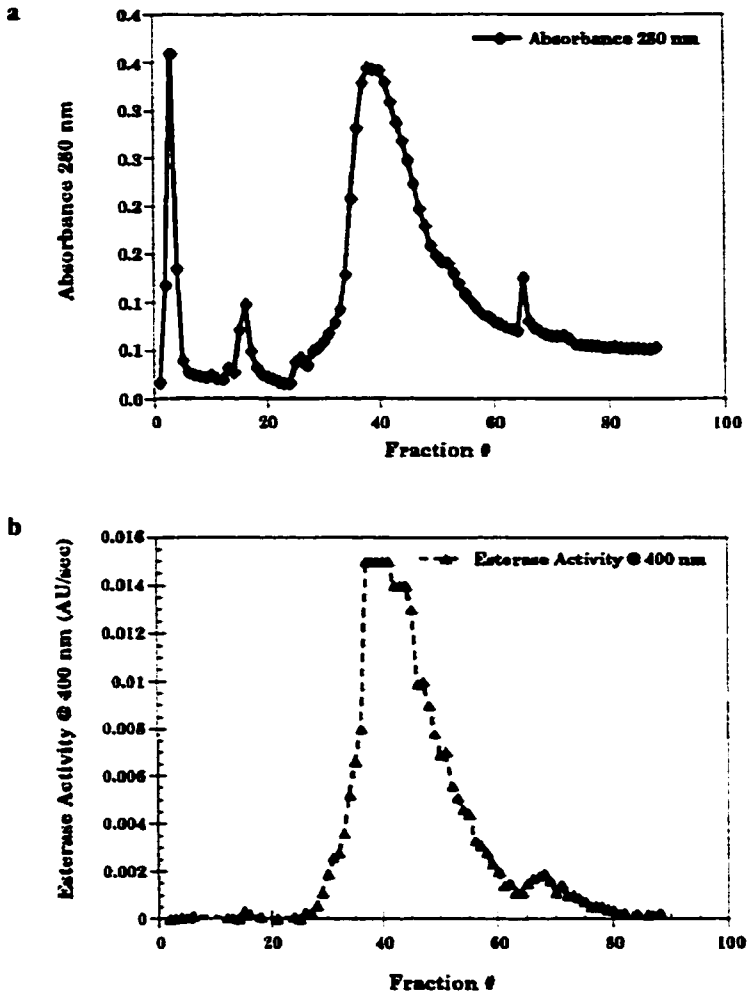


Figure 6-2B. (a) Protein elution profile as measured by absorbance at 280 nm. A switch from the initial elution buffer to 8 mM taurocholate buffer was made at fraction number (F) 13, followed by an 8-40 mM taurocholate buffer at F25. (b) Esterase activity elution profile for BCE (using in-house-prepared cholate-derivatized Sepharose 4B affinity column) as determined using *p*-NPB.

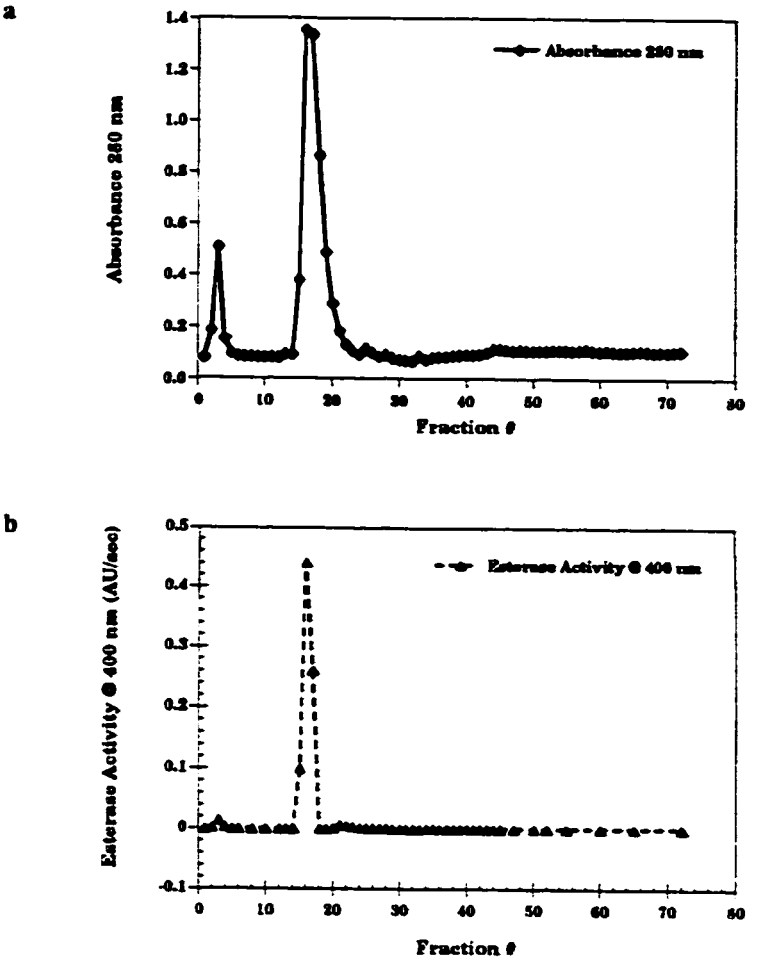


Figure 6-2C. (a) Protein elution profile as measured by absorbance at 280 nm. A switch from the initial elution buffer to 8 mM taurocholate buffer was made at fraction number (F) 13, followed by an 8-40 mM taurocholate buffer at F33. (b) Esterase activity elution profile for BCE (using Sigma cholate-derivatized agarose affinity column) as determined using *p*-NPB.

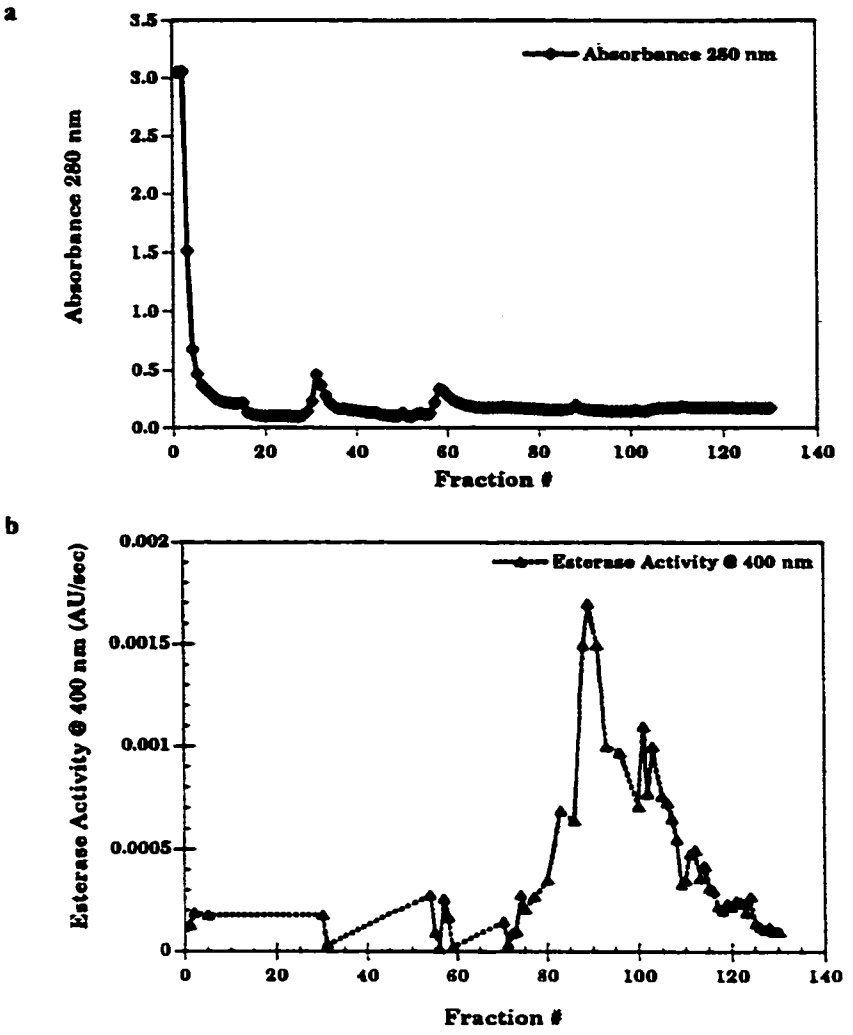


Figure 6-2D. (a) Protein elution profile as measured by absorbance at 280 nm. A switch from the initial elution buffer to 8 mM taurocholate buffer was made at fraction number (F) 24, followed by an 8-40 mM taurocholate buffer at F52. (b) Esterase activity elution profile for human milk BAL (using in-house-prepared cholate-derivatized Sepharose 4B affinity column) as determined using *p*-NPB.

esterase activity in this peak indicates that the column retains all of the cholesterol esterase that is loaded. To evaluate the loading capacity of the column, incrementally larger quantities of PCE were loaded, until esterase activity appeared in the wash-through, indicating that the enzyme binding sites were saturated. The loading capacity for the above-specified column (7 mL of packed gel) was approximately 600 mg crude material (82 activity units). Optimal loading, for purification purposes, was under 100 mg crude material. For larger amounts of material, elution with each buffer must be continued for a longer period. An additional peak without esterase activity was observed to elute shortly after the 8 mM taurocholate wash was started (approximately 60-70 mL). Concentration and subsequent electrophoresis of these combined fractions showed a band with a molecular mass of approximately 45 kDa. It showed protease activity when tested with *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide. This protease was not characterized further. It is noteworthy, however, that it has an affinity for the cholate-Sepharose, albeit weak. It is doubtful that this affinity is due to nonspecific binding, as elution of the protease is not observed if the initial wash is prolonged.

From the esterase activity profile it can be seen that the enzyme begins to elute after passage of a volume of approximately 100 mL, starting at a taurocholate concentration of approximately 25 mM. The elution of the porcine enzyme is quite sharp, with esterase activity dropping close to background levels within 40 mL. The bovine enzyme elutes over a larger volume, with slight esterase activity still present after elution of 60 mL.

Figure 6-3(a) shows an SDS-PAGE (gradient gel) of affinity column-purified BCE, obtained from a commercially available crude preparation, illustrating the high

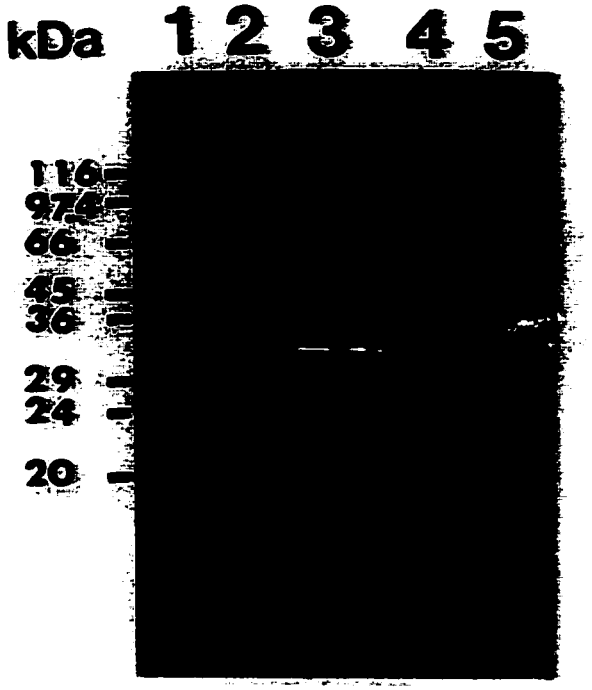


Figure 6-3(a). PhastGel™ SDS-PAGE (10-15 % gradient gel) of commercial (lane 4) and affinity purified (lanes 2 and 3) BCE. The commercial and affinity purified samples were heated at 100°C for 5 min. in 0.4% and 2.5% SDS-sample buffer, respectively. Lanes 1 and 5: protein molecular weight standards.

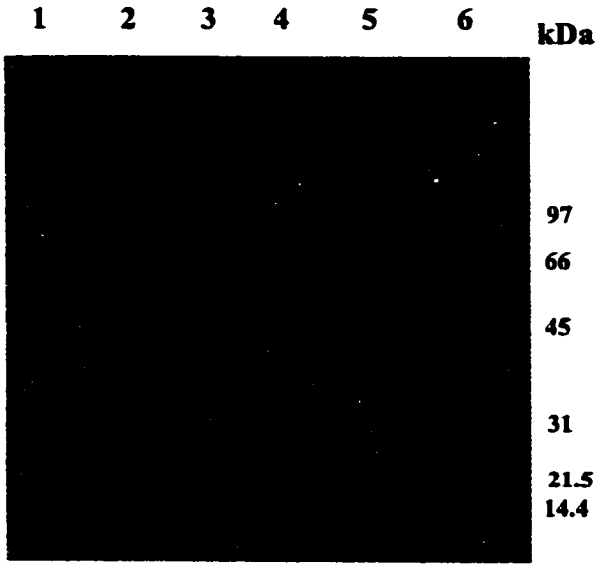


Figure 6-3(b). PhastGel™ SDS-PAGE (12.5% homogeneous gel) of partially affinity-purified commercial (lane 4) and affinity-purified (lanes 1-3, 5) PCE; lane 6, protein molecular weight standards.

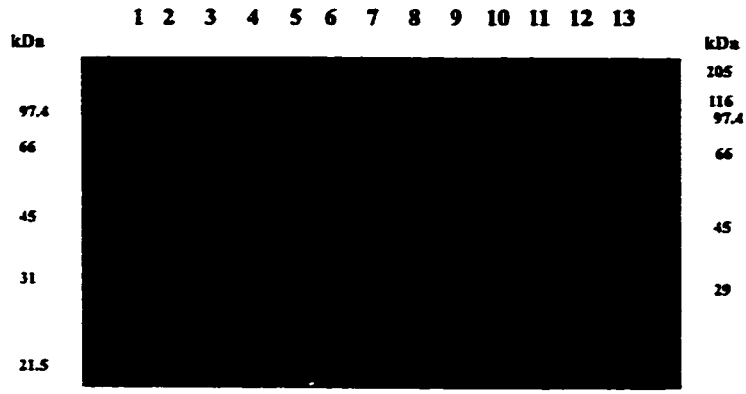


Figure 6-3(c). BIO-RAD SDS-PAGE slab gel (12% homogeneous) of crude and affinity-purified PCE. Lanes 1, 2, low molecular weight marker proteins; lanes 3, 4, concentrated affinity-purified PCE (dimer is evident); lanes 5-9, less concentrated affinity-purified PCE; lane 10, high molecular weight marker proteins; lane 11, crude PCE; lanes 12, 13, affinity-purified PCE.

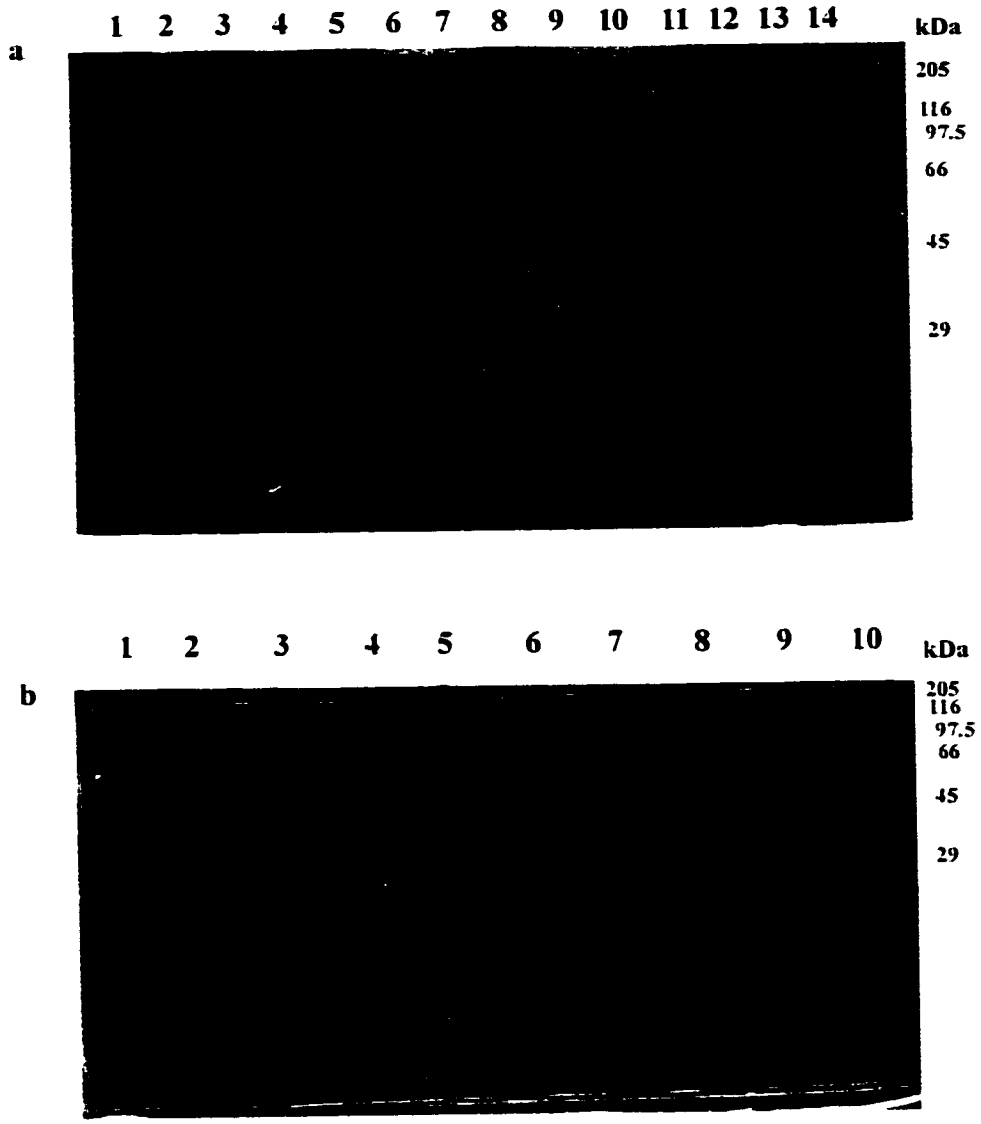


Figure 6-4. BIO-RAD slab SDS-PAGE gels (12% homogeneous) of crude and affinity-purified human milk bile salt-activated lipase (BAL).

Bands in each lane (#) correspond to proteins in the following samples. For gel (a), (1), (12), low molecular weight markers (LMWM); (2) crude "red" milk supernatant; (3) fractions 1-3 washthrough milk proteins; (4)-(6) fractions 4-24, 25-52, and 53-72, milk proteins; (7)-(11) fractions 73-87, 88-94, 95-103, 104-127, 128-130, pure BAL; (13), (14), high molecular weight markers (HMWM). For gel (b), (1), empty; (2), (9), HMWM; (3)-(6), affinity-purified human milk BAL; (7), (8), crude human milk BAL; (10), LMWM.

Gel (a) illustrates the progressive increase in purity of eluted BAL with an increase in the taurocholate concentration in the eluant. A single band for pure BAL corresponded to the fractions containing esterase activity (see esterase activity elution profile in Figure 1D). Bands of affinity-purified human milk BAL fractions were not very intense due to the dilute nature of the samples in (a), but those in (b) are more intense. Gel (b) illustrates the efficacy of the cholate-affinity column.

specificity of the affinity column for cholesterol esterase (apparent MW 65 kDa; SDS PAGE performed on 12 % homogeneous gels gave an apparent MW of 58 kDa). The faintness of the band at 65 kDa in the crude enzyme preparation (lane 4) likely is a result of proteolysis of the BCE by contaminating proteolytic enzymes.

The high specificity of the in-house-prepared cholate-derivatized affinity column was also illustrated for the purification of PCE. Figure 6-3(b) shows an SDS-PAGE (gradient gel) of affinity column-purified PCE, obtained from a commercially available crude preparation. Two distinct monomeric forms were observed at approximately 70 and 80 kDa. Porcine CE has been found to exist in structurally distinct monomeric and dimeric forms (i.e., monomer, 74±4 kDa; dimeric subunit, 83±4 kDa; Rudd *et al.*, 1987). The two distinct monomeric forms were observed in fraction samples that were dilute (Figure 6-3(b)); however, in more concentrated samples (i.e., insufficient SDS present in sample buffer; taurocholate-induced dimerization; see Part B), a dimeric form (approximately 170 kDa) was observed (Figure 6-3(c), lanes 3 and 4). The literature value for a PCE dimer is 167 kDa (Rudd *et al.*, 1987).

The purification factors, as well as yields for the purification of porcine and bovine cholesterol esterase are listed in Table 6-1. The purification factor for PCE is 40-fold with 75 % of total activity recovered, which is an excellent result for a single step procedure. The corresponding purification for BCE was less efficient but still impressive. The use of taurocholate in the eluent was crucial to the success of the method. Replacement of taurocholate with cholate did not result in a satisfactory purification. A potential advantage of taurocholate is that, being the conjugate base of a strong acid, it will have little effect on the pH of the buffer, whereas cholate, being the

conjugate base of a weak acid, will increase the pH when its concentration is comparable to that of the buffer. Cholate-induced shifts in the pH of the buffer could, for example, cause nonspecific elution of proteins from the column.

The choice of immobilized support also appears to be important. In our hands, we had success with EAH Sepharose 4B (4% beaded agarose, 10 carbon spacer arm) but had no success with a commercially available, cholate-derivatized, cross-linked 4% beaded agarose (9 carbon spacer arm; Sigma) using either cholate or taurocholate as the eluent. In the latter case, for example, BCE eluted earlier with 8mM taurocholate along with contaminating protease activity. The cause of the difference in behaviour in the Sepharose and the commercial material is not obvious. It is unlikely that it arises from differences in the number of carbons in the spacer arm. The larger mesh size of the commercial resin may have resulted in a poorer separation.

Table 6-1. Specific activities, purification factors, and activity yields for porcine and bovine cholesterol esterases before and after purification by affinity chromatography

Enzyme	Specific activity (crude) ($\mu Ms^{-1}mg^{-1}$)	Specific activity (pure) ($\mu Ms^{-1}mg^{-1}$)	Purification (fold)	Total activity loaded (μMs^{-1})	Total activity recovered (μMs^{-1})	Activity Yield (%)
PCE	0.2	8	40	4	3	75
BCE	0.6	5	8	7	2	29

PART B. Taurocholate-Induced Dimerization of Bovine Cholesterol Esterase in Sodium Dodecylsulfate

2B. Material and methods

Purified bovine CE was purchased from Genzyme Inc. [^{14}C]Taurocholic acid was obtained from *NEN* Research Products (specific activity: 46.4 mCi/mmol). All other chemicals, reagents and solvents were high purity preparations obtained from commercial sources.

2.1. Electrophoresis

The Pharmacia PhastGel™ system and the BioRad vertical slab gel apparatus were employed. Electrophoresis was conducted primarily on the former system and the latter gel apparatus was employed for preparative gels as well as to verify the results obtained on the PhastGel™ system. In the BioRad vertical slab gel case, stock solutions for stacking and resolving gel preparations were made according to Hames (1990) for discontinuous SDS-PAGE. All gels were stained with Coomassie blue.

2.2. Analytical SDS-PAGE Sample Preparation

Sample buffers were 5x concentrated using the original recipe obtained from Bollag and Edelstein (1991). Total sample volumes were 20 μL. Samples were prepared in such a manner as to test each variable independently, while keeping others constant. The variables were: (i) CE concentration, (ii) SDS concentration, (iii) heating time at 100°C, (iv) β-mercaptoethanol concentration, (v) taurocholate concentration, (vi) other bile salts and (vii) urea. Two molecular mass standard protein mixtures were employed. Standard I contained proteins with molecular masses: 97.4, 66.2, 45, 31, 21.5, and

14.4 kDa. Standard II contained proteins with molecular masses: 116, 97.4, 66, 45, 36, 29, 24, 20, and 14.2 kDa.

2.2.1 Sample Buffers

The sample buffer employed for preparation of protein samples for SDS-PAGE usually contains 0.4 % SDS (Hames, 1990). This sample buffer was employed in the determination of the effect of bovine CE concentration on the amount of observable dimer on SDS-PAGE. For convenience, a sample buffer containing no SDS was used in the experiments on the effect of SDS concentration on dimer formation, with the appropriate amounts of SDS stock solutions added to each sample separately. In a similar manner, the study on the effect of β-mercaptoethanol on the amount of observable dimer at specific SDS concentrations employed a stock sample buffer solution absent in this reducing reagent, with the appropriate amounts of β-mercaptoethanol stock solutions added to each sample separately. Urea-containing sample buffer was 6 M in urea and was also absent in SDS for the convenience of creating the appropriate SDS concentration in protein samples by adding the appropriate amounts of SDS stock solutions. Sample buffer recipes are given below for 10 mL total volume.

5x–2% SDS-sample buffer: 0.6 mL 1 M TrisHCl, pH 6.8 (60 mM); 2.5 mL 100% glycerol (25%); 2.0 mL 10% SDS (2%); 0.5 mL β-mercaptoethanol (0.717 M); 1 mL 1% bromophenol blue (0.1%); and 0.9 mL distilled H_2O.

5x–0% SDS-sample buffer: 0.6 mL 1 M TrisHCl, pH 6.8 (60 mM); 2.5 mL 100% glycerol (25%); 0.5 mL β-mercaptoethanol (0.717 M); 1 mL 1% bromophenol blue (0.1%); and 5.4 mL distilled H_2O.

5x–0% SDS-0% β–ME-sample buffer: 0.6 mL 1 M TrisHCl, pH 6.8 (60 mM); 2.5 mL 100% glycerol (25%); 1 mL 1% bromophenol blue (0.1%); and 5.9 mL distilled H_2O.

5x–6 M UREA-0% SDS-sample buffer: 0.6 mL 1 M TrisHCl, pH 6.8 (60 mM); 2.5 mL 100% glycerol (25%); 0.5 mL β-mercaptoethanol (0.717 M); 1 mL 1% bromophenol blue (0.1%); 3.6036 g 99% ultra pure urea; and distilled H_2O up to 10 mL.

2.2.2 Effect of SDS concentration on the amount of observable dimer

Each sample contained a total sample volume of 20 μL comprised of 160 μg of protein (commercially purified bovine CE). To 8 μL of a 20 mg/mL bovine CE stock solution in distilled water (contained in an eppendorf tube) were added the appropriate volumes of distilled H_2O and SDS stock solution prior to the addition of 5x–0% SDS-sample buffer (which contained no SDS). The appropriate amount of SDS was added to the protein sample for convenience (i.e., instead of preparing fourteen different SDS-sample buffers). The concentrations of SDS in each sample ranged from 0, 0.05, 0.1, 0.2, 0.4, 0.5, 0.8, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, and 2.5 %. In microgram quantities, these percentages correspond to 0, 10, 20, 40, 80, 100, 160, 200, 250, 300, 350, 400, 450, and 500 μg SDS in a total sample volume of 20 μL, respectively. Samples were boiled for 5 minutes prior to being electrophoresed on the Pharmacia PhastGel™ system. SDS-PAGE 10-15% gradient PhastGels™ were employed.

2.2.3 Effect of bovine cholesterol esterase concentration on the amount of observable dimer

Each sample of 20 μL total volume contained 0.4 % SDS, while sample protein concentration was varied from 10, 8, 6, 4, 2 and 1 mg/mL (or μg/μL). To 10, 8, 6, 4, 2 or

1 μL of 20 mg/mL bovine CE contained in separate eppendorf tubes were added 6, 8, 10, 12, 14, or 15 μL of distilled H_2O, respectively, and 4 μL of 5x–2% SDS-sample buffer. Each sample was boiled for 5 minutes prior to conducting SDS-PAGE on the Pharmacia PhastGel™ system using 10-15% gradient SDS-PAGE PhastGels™.

2.2.4 Effect of boiling time in SDS-sample buffer on the amount of observable dimer

Each sample of 20 μL total volume contained 8 mg/mL bovine CE and 0.2 % SDS. The latter amount of SDS was chosen because it gave the highest amount of observable dimer upon boiling for 5 minutes compared to other %SDS concentrations used (Figure 6-6C(i), lane 5). Each of the seven samples was prepared in the following manner. To 8 μL of 20 mg/mL bovine CE contained in an eppendorf tube were added 4 μL of distilled H_2O, 4 μL of 1% SDS stock solution (40 μg SDS in a total volume of 20 μL) and 4 μL of 5x–0 %SDS-sample buffer. One of the seven samples was not boiled, while the others were boiled for 30 sec., 1, 2, 5, 10 and 15 min, respectively, before SDS-PAGE was conducted on the PhastGel™ system using 10-15% gradient SDS-PAGE PhastGels™.

2.2.5 Effect of taurocholate concentration on the amount of observable dimer

Each of seven samples of 20 μL total volume contained 8 mg/mL bovine CE and 2.5% SDS. This percentage of SDS was employed because it is the minimum concentration of SDS in a sample where only monomer can be observed upon boiling for 5 minutes; this eliminated the effect of SDS on the amount of observable dimer so that the effect of taurocholate concentration on the latter could be studied. Each sample was prepared in the following manner. To 8 μL of 20 mg/mL bovine CE contained in an

eppendorf tube were added 3 μL (or 5 μL) of the appropriate concentration of taurocholate stock solution [3 μL of 0, 66.67, 200, or 266.67 mM taurocholate stock solutions to respectively make 0, 10, 20, 30, and 40 mM taurocholate in the 20 μL total sample volume; or 5 μL of 200 or 240 mM taurocholate stock solutions to respectively make 50 and 60 mM taurocholate-containing samples], followed by 5 μL of 10% SDS stock solution (or 3 μL of 16.67% SDS) and 4 μL of 5x–0% SDS-sample buffer. Samples were boiled for 5 minutes prior to being run on the PhastGel™ system using 10-15% gradient SDS-PAGE PhastGels™.

2.2.6 Effect of 6 M urea in the sample buffer on monomer and dimer electrophoretic properties on SDS-PAGE

Each 20 μL total volume sample contained 8 mg/mL bovine CE and 0.2% SDS. Six samples contained SDS-sample buffer, while another six were prepared with 6 M urea present in the SDS-sample buffer. Each of the twelve samples was prepared in the following manner. To 8 μL of 20 mg/mL bovine CE contained in an eppendorf tube were added 4 μL of distilled H_2O, 4 μL of 1% SDS stock solution (40 μg SDS in a total volume of 20 μL) and either 4 μL of 5x–0% SDS-sample buffer or 4 μL of 5x-6 M urea–0% SDS-sample buffer. One of the six samples containing SDS but not urea was not boiled, while the others were boiled for 5, 15, 25, 35, and 45 minutes, respectively. The six urea-SDS-containing samples were treated in a similar fashion. A comparison of the electrophoretic mobilities of the monomer and dimer bands was conducted by running similarly treated (i.e., with respect to boiling time) SDS-containing samples beside urea-SDS-containing samples on SDS-PAGE employing the PhastGel™ apparatus and 7.5% homogeneous PhastGels™.

2.3. Isolation of the CE monomer and CE dimer from SDS-PAGE gels

Purified bovine CE (2.5 mg) was loaded onto a preparative BioRad slab gel apparatus. The protein sample contained 8 mg/mL bovine CE and 0.2% SDS and was boiled for 10 minutes prior to loading the gel. After electrophoresis, cutting and staining two longitudinal "marker" strips from each side of the slab gel located the bands corresponding to the monomer and dimer. The regions of the unstained gel containing the monomeric and dimeric CE were excised and eluted according to the procedure described by Hames (1990). Monomer and dimer proteins were eluted by diffusion into buffer from their respective isolated gel strips in the following manner. Each strip of isolated monomer or dimer was placed in a large screw-cap vial and macerated with a glass rod. Elution was carried out with approximately three gel volumes of buffer (50 mM Tris, pH 8.6 in 0.1% SDS) overnight with mixing. The samples were then placed on a rotating flatbed that was heated at approximately 50°C for overnight (18 hours). The gel mass was pelleted out via centrifugation and a repeated extraction was carried out the next day. The combined extracts of monomer and dimer were dialyzed against distilled water, lyophilized, and analyzed for stability.

2.4 Characterization of the interaction between taurocholate and bovine CE

To determine whether taurocholate is covalently associated with the dimeric species of CE, protein samples were incubated with radiolabeled taurocholate prior to being electrophoresed in the following manner. Two samples were prepared in eppendorf tubes, each containing: 0.00432 μmol of [24-^{14}C]-taurocholate (the 1:3 methanol:ethanol solvent in which the taurocholate was dissolved and removed from the sample using a speed vacuum); 8 μL of 20 mg/mL bovine CE for a concentration of

8 μg/μL in a total volume of 20 μL; 4 μL of distilled H_2O; 4 μL of 1% SDS stock solution (i.e., 40 μg SDS in a total volume of 20 μL; 0.2% SDS in sample); and 4 μL of 5x-0% SDS-sample buffer. The samples were boiled for 0 or 15 minutes and each sample was loaded onto two Pharmacia PhastGels™ and onto seven wells per gel (each lane loads about 1 μL of sample onto the gel). Sample strips from one gel were prepared for liquid scintillation counting, while the second gel of the same sample was employed for autoradiography. The gels were run on the Pharmacia PhastGel™ system but were not stained and destained as usual in order to avoid the diffusion of the radiolabeled taurocholate from the gel. Instead, outer strips, one containing a marker lane and a sample lane, and the other containing just the sample lane, were cut out and stained with Coomassie blue and then destained. These stained "marker" strips were employed to identify the relative positions of the monomer and dimer bands for the inner untreated gel. Horizontal strips of monomer and dimer bands (as well as other regions of the unstained gel) were scraped off one of each of the PhastGels™ that contained samples that were boiled for 0 or 15 minutes. Individual sample strips were placed into large (approximately 10 mL) screw-capped test tubes and allowed to dry overnight at room temperature prior to solubilizing the bisacrylamide-crosslinked polyacrylamide gels using hydrogen peroxide. Depending on the amount of gel required to be solubilized, multiples of 0.25 mL of 30% (w/v) H_2O_2 were added to each screw-capped test tube. Each sample tube was tightly capped to prevent the escape of radioactive $^{14}CO_2$ and was incubated on an angle at 50 °C overnight to allow solubilization. Sample tubes were cooled to room temperature before opening and to each were added 5 mL (or multiples of, to keep the

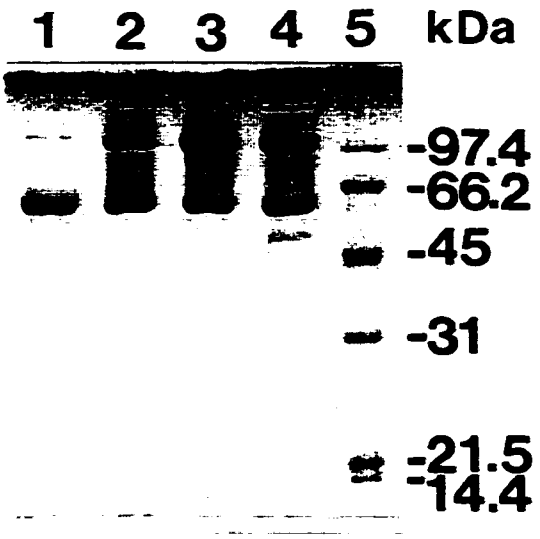


Figure 6-5. Comparison of affinity-purified and commercial bovine CE.

SDS-PAGE 12% homogeneous gel of purified bovine CE obtained from Genzyme Inc. (lane 1), affinity-purified bovine CE (lanes 2-4) and molecular weight standards (lane 5). Samples containing 0.4% SDS were heated for 5 minutes.

level of quenching constant) of scintillation cocktail. The resulting counting mixtures were completely clear.

3B. Results

In an attempt to remove trace impurities from bovine CE, the enzyme was bound to a cholate-derivatized Sepharose affinity column (Part A; Moore *et al.*, 1996) and eluted with taurocholate. SDS-PAGE of the starting material showed one major band with an apparent molecular mass of 58 kDa (Figure 6-5, lane 1) and several minor impurities. The protein eluted from the affinity column showed an additional major band with an apparent molecular mass of 116 kDa (Figure 6-5, lanes 2-4), giving the impression that the purification procedure had actually increased the heterogeneity of the protein. A small amount of the 116 kDa protein is also visible in the commercial enzyme (Figure 6-5, lane 1) but unlike the affinity-purified enzyme it is a minor component. Proteolysis of the 116 kDa component was not involved as treatment with DFP to inhibit serine proteases, EDTA to inhibit metalloproteases or lowering the pH of the sample buffer to 4 to inhibit alkaline proteases had no effect. The mass relationship between the two major bands strongly suggests that they represent monomeric and dimeric forms of the same polypeptide chain. On gradient SDS-PAGE gels the corresponding apparent molecular masses were 65 kDa and 130 kDa (Figure 6-6) but the values obtained from the homogeneous gels are likely the more accurate. To avoid confusion in describing and discussing the results obtained, these protein bands on the gradient gels are also designated as 58 kDa (monomer) and 116 kDa (dimer).

The apparent amount of the 116 kDa band relative to the 58 kDa band increased with CE concentration when incubated in the SDS-sample buffer at 100°C for 5 minutes

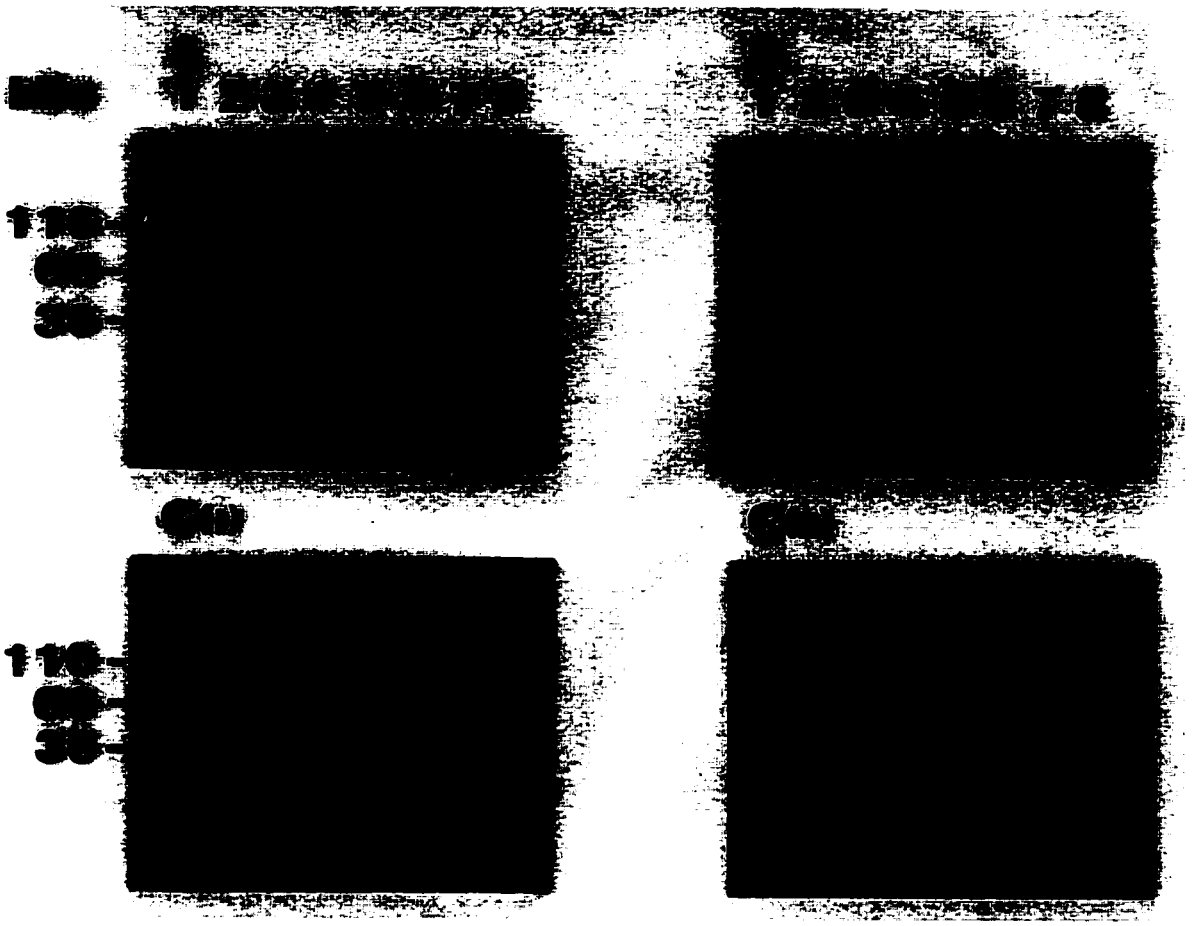


Figure 6-6. Effect of conditions on sample preparation.

Samples were run on SDS-PAGE 10-15% gradient PhastGel™ gels.

(A) CE concentration. CE samples containing 0.4% SDS were heated for 5 min at 100°C. Samples in lanes 2-7 contained the following concentrations of CE: 1, 2, 4, 6, 8, and 10 mg/mL.

(B) Heating time. Samples of CE (8 mg/ml) containing 0.2% SDS were heated at 100°C for the following times in lanes 2-8: 0, 0.5, 1, 2, 5, 10 and 15 min, respectively.

(C) SDS concentration. CE (8 mg/ml) samples were heated for 5 min at 100°C. Samples in lanes 2-8 of C(i) contained the following concentrations of SDS: 0, 0.05, 0.1, 0.2, 0.4, 0.5, and 0.8%, respectively. Samples in lanes 1-7 of C(ii) contained the following concentrations of SDS: 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, and 2.5%, respectively.

(Figure 6-6A). Longer heating times in the SDS-sample buffer also increased the amount of dimer relative to monomer (Figure 6-6B). These results are consistent with the association of monomeric subunits. At very long heating times, high molecular weight aggregates became apparent at the top of the gel (Figure 6-6B, lane 8). However, only the monomer was present when CE was not heated in SDS-sample buffer at 100°C (Figure 6-6B, lane 2). The relative amount of dimer decreased as the SDS concentration in the sample increased and no dimer was observed at concentrations of SDS greater than 2.25% (Figure 6-6C(i) and 6-6C(ii)). This observation suggests that another factor whose effect can be eliminated at high SDS concentration is involved in promoting dimer formation.

Taurocholate was used to elute CE from the affinity column and its presence may be responsible for promoting the formation of the 116 kDa dimer band on SDS-PAGE (Figure 6-5, lanes 2-4). Indeed, addition of taurocholate to CE had the effect of producing dimer under conditions where only monomer was previously observed (Figure 6-7A). Cholate, deoxycholate and glycochenodeoxycholate also promoted dimer formation but at 20 mM concentration taurocholate appeared to have the greatest affect (Figure 6-7B). When [^{14}C]taurocholate was added to the CE sample, no radioactivity was found to be associated with the dimer, indicating that the dimer does not contain either covalently or noncovalently associated taurocholate. Similarly, addition of up to 2 M β–mercaptoethanol and 100 mM hydrazine had no effect on the dimerization, indicating that no disulfide bridges or amide linkages are formed between the polypeptide chains in the dimerization process.

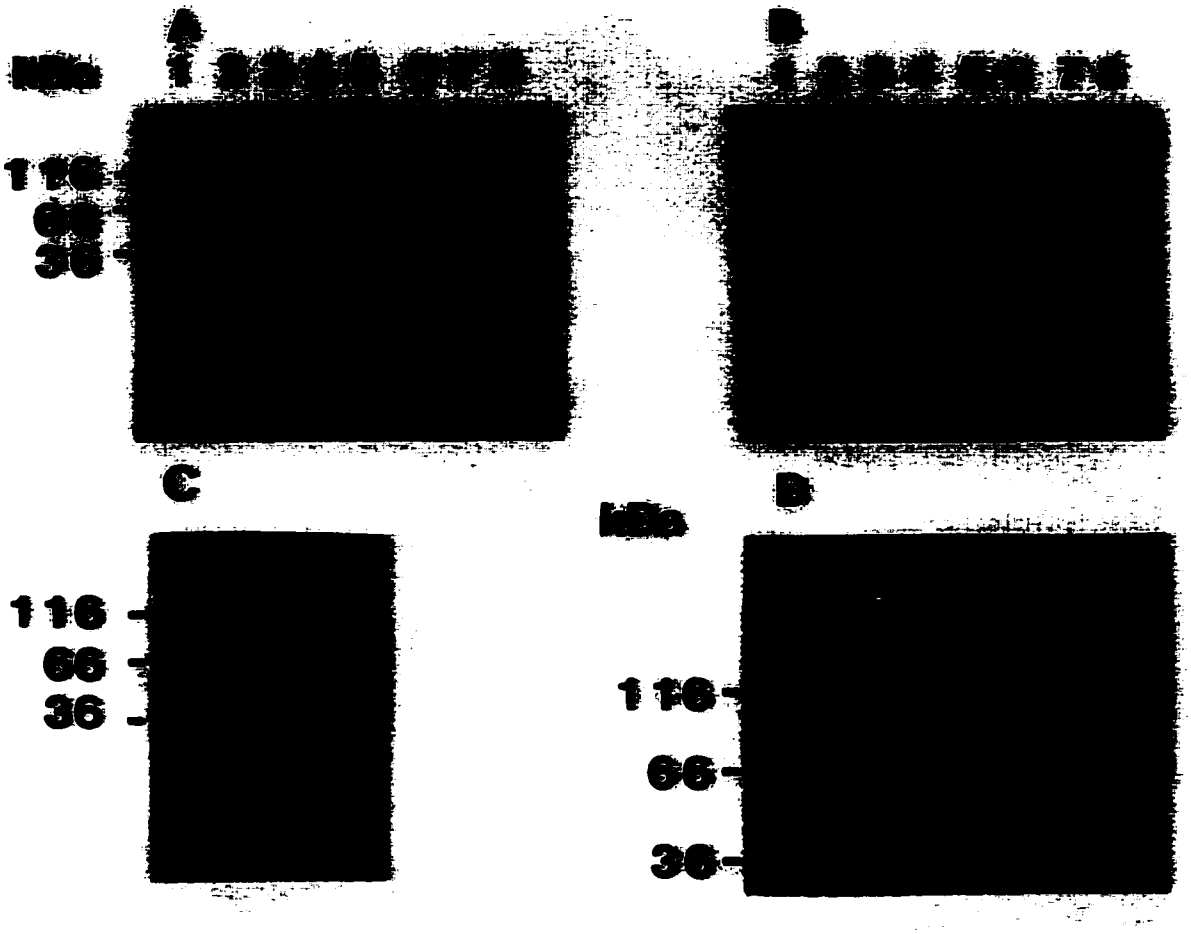


Figure 6-7. Factors affecting the association of CE subunits.
Unless otherwise stated, samples were run on SDS-PAGE 10-15% gradient PhastGel™ gels. **(A) Taurocholate concentration.** Samples of CE (8 mg/ml) containing 2.5% SDS were heated for 5 min at 100°C. Taurocholate concentrations in lanes 2-8 were: 0, 10, 20, 30, 40, 50 and 60 mM, respectively. **(B) Other bile salts.** CE (8 mg/mL) samples in 2.5% SDS were heated for 5 min. at 100°C in the absence (lanes 2 and 7) or presence of 20 mM bile salt (taurocholate, cholate, deoxycholate and glycochenodeoxycholate; lanes 3-6, respectively). **(C) Isolated monomeric and dimeric forms of CE.** Lanes 1 and 2 contain the isolated 116 kDa protein, while lanes 4 and 5 contain the isolated 58 kDa protein. Lanes 1 and 4, and 2 and 5, are comparable, respectively. Samples in lanes 1 and 4 contained 2.5% SDS and were heated for 5 minutes; while those in lanes 2 and 5 contained 0.2% SDS and were not heated. Lane 3 shows the molecular weight protein standards. **(D) Urea.** SDS-PAGE 7.5% homogeneous PhastGel™ gel of bovine CE (8 mg/mL) samples in 0.2% SDS. Lanes 2, 4 and 6 contain samples prepared with SDS-sample buffer and heated for 25, 35, or 45 minutes, respectively. Lanes 3, 5 and 7 contain samples prepared with 6 M urea-SDS-sample buffer and heated for 25, 35, or 45 minutes, respectively. Lane 8 contains an unheated CE sample prepared with SDS-sample buffer. Lane 1 shows the molecular weight marker proteins.

The apparent monomer-dimer relationship between the protein in the 58 kDa and 116 kDa bands was investigated further by isolating and concentrating the protein from the respective bands. When the 116 kDa protein was rerun on SDS-PAGE containing 2.5% SDS, a 58 kDa band predominated (Figure 6-7C, lane 1). With 0.2% SDS, more of the 116 kDa band was present but the 58 kDa band was still the major component (Figure 6-7C, lane 2). The isolated 58 kDa protein remained predominantly as a 58 kDa band but a minor 116 kDa component was generated (Figure 6-7C, lanes 4 and 5).

Addition of 6M urea to the SDS-sample buffer reduced the electrophoretic mobility of the monomer and dimer bands (Figure 6-7D, lanes 3, 5 and 7). In the absence of urea, the tendency of CE to form higher molecular weight aggregates on long heating times in SDS-sample buffer became more pronounced (Figure 6-7D, lanes 2, 4 and 6). The presence of urea prevented this aggregation but did not greatly affect the association of monomers to form dimer.

3B. Discussion

Heating of proteins in sample buffer prior to performing SDS-PAGE is carried out to unfold the protein and to ensure that sufficient SDS is bound to prevent association of polypeptide chains. It was therefore unexpected that this procedure would promote the association of subunits to form a dimer of CE. Only the presence of monomer on SDS-PAGE has previously been reported with bovine CE (Cox *et al.*, 1990) but unfortunately insufficient experimental detail was given to make a direct comparison with our results. The monomer-dimer relationship between the 58 kDa and 116 kDa proteins was confirmed by the observation that the isolated 116 kDa protein reverted to the 58 kDa protein when rerun on SDS-PAGE. Protein concentration, SDS concentration

and heating time in the sample buffer influenced the amount of dimerization observed, but the major effect was due to the presence of taurocholate (Figure 6-7A).

In addition to dimer formation, high molecular weight aggregates were observed with long heating times and high taurocholate concentrations (Figure 6-6B, lane 8; Figure 6-7A, lanes 5-8). The aggregation appeared to be similar to that observed with membrane proteins which often strongly associate to give rise to aggregates when subjected to SDS-PAGE (Clarke *et al.*, 1985; Larson *et al.*, 1983; Stocker *et al.*, 1983; Teather *et al.*, 1978; Talkington *et al.*, 1992; Gallagher *et al.*, 1987). As the dimer of CE appeared before any aggregate formation, the high molecular weight aggregates of CE were likely formed by the association of dimers. It appears that the same factors that promote dimer formation also promote the formation of higher molecular weight aggregates but under most conditions the dimer appeared to be the predominant associated form.

The observations that high concentrations of SDS prevent dimer formation and that the dimer readily reverts to the monomer indicate that hydrophobic interactions between monomeric subunits are responsible for dimer formation. The results obtained show that cholate, deoxycholate and glycochenodeoxycholate also promote dimerization but appeared to be less effective than taurocholate (Figure 6-7B). It is therefore likely that a specific noncovalent interaction of taurocholate with CE is responsible for the dimerization observed.

It is well established that there is an obligatory requirement for 3α,7α,12α-trihydroxycholanates, i.e. cholate and its conjugates, for the activation of CE (Calame *et al.*, 1975; Erlanson, 1975; Hyun *et al.*, 1969, 1971, 1972; Jacobson *et al.*, 1989;

Lombardo *et al.*, 1978; Meyer, 1989; Rudd *et al.*, 1987; Vahouny and Brecher, 1968; Vahouny *et al.*, 1965). However, their effect on association appears to vary with the source of the CE. In the case of CE from rat pancreatic juice, the binding of bile salt to the monomeric unit (Calame *et al.*, 1975; Hyun *et al.*, 1971, 1972) appears to induce association to form hexamers. On the other hand, human pancreatic CE (Lombardo *et al.*, 1978; Meyer, 1989) forms tetramers in the presence of bile salts. Porcine CE (Rudd *et al.*, 1987) was found to exist in structurally distinct monomeric and dimeric forms and bile salt was found to have no effect on association. Bovine pancreatic CE (Cox *et al.*, 1990) has been reported to exist in two distinct polypeptide forms but no details were given on the state of association of these forms in the presence of bile salts. The present study demonstrated that taurocholate mediates the dimerization of bovine CE under denaturing conditions. However, there is no evidence that a dimer exists under physiological conditions and therefore, no significance of this highly unusual dimerization phenomenon can be attributed to CE function. The fact that such a specific interaction does occur under denaturating conditions suggests that under nondenaturing conditions bile salts induce oligomerization of bovine CE. Indeed, under denaturing conditions with high concentrations of taurocholate, small amounts of higher molecular weight aggregates were observed (Figure 6-7A).

It is not obvious why taurocholate or other bile salts promote the association of monomers to form dimers on heating in SDS. However, the observation that the presence of urea in the sample buffer reduced the electrophoretic mobility of the monomer and dimer indicates that there are strong intrachain hydrophobic interactions which are not readily disrupted by SDS leaving some regions of the protein partially folded. It is

possible that the regions involved in these intrachain hydrophobic interactions gradually become exposed during the time the protein is heated in the sample buffer. As the protein sample cools, intrachain interactions can be reformed, interchain interactions can be formed or SDS can be bound. High concentrations of CE would be expected to increase the likelihood of interchain interactions to form dimer and this is consistent with the results obtained in the present study. Taurocholate promotes the association of subunits and therefore its presence would increase the effective concentration of subunits for dimer formation. At high SDS concentrations, formation of the monomer is favored by increased binding of the SDS to the newly exposed hydrophobic regions. When the CE samples containing taurocholate are not heated, the SDS unfolds the subunits destroying the quaternary structure of the protein but small regions containing intrachain hydrophobic interactions are not completely unfolded. These regions involved in dimer formation are therefore not exposed so only the monomeric unit is observed on SDS-PAGE when the sample is not heated.

The present investigation shows that the conditions of sample preparation prior to performing SDS-PAGE can have unexpected effects on the apparent homogeneity of a protein. It is common practice to heat proteins for long periods of time in SDS in order to break strong hydrophobic associations but, as observed in the present study, this can cause association of polypeptides if other factors that promote interchain interactions are present. In cases where multiple bands are observed on SDS-PAGE, care should therefore be taken to verify that the bands observed represent the individual monomeric components of the protein and not associated forms.

3B.1 Prospective Studies for Cholesterol Esterase: Use of Low Water Reverse Micellar Systems to Probe Structure-Function Relationships

3B.1.2 Characteristics of Reverse Micelles

Reverse micelles are spherical water droplets surrounded by a monolayer of closely packed surfactant molecules dispersed in a solvent of low polarity (Luisi *et al.*, 1988). Surfactant molecules have their polar head groups in contact with and hydrated by water molecules that lie in the interior of the micelle, while their hydrophobic tails are in contact with the surrounding organic solvent. Such systems are thermodynamically stable due to the presence of the interfacial surfactant layer that prevents unfavourable direct contact between water and organic solvent.

At a fixed concentration of surfactant, the dimension of the water pool inside the micelles varies in proportion to the amount of water that is introduced (i.e., the ratio of water to surfactant molecules ($[H_2O]/[surfactant]$), W_0, increases: θ = % H_2O, v/v). Thermodynamically stable reverse micelles have dimensions in the range from 0 to 10 nm (Zulauf and Eicke, 1979).

Reverse micellar size remains constant when water and surfactant concentrations are varied proportionally, but the concentration of reverse micelles changes (Luisi *et al.*, 1988). At typical water and surfactant concentrations, the concentration of reverse micelles in an organic solvent may range from 0.1 to 10 mM, (typical enzyme concentration is nM or μM; substrate concentration is μM or mM).

3B.1.3 Properties of Water in Reverse Micelles

The properties and state of water in the interior of reverse micelles formed with dioctyl sodium sulfosuccinate (AOT), phospholipids or cetyltrimethylammonium bromide (CTAB) (Darszon and Shoshani, 1992; Kernen *et al.*, 1997) have been studied

by spectroscopic techniques such as NMR and ESR spin labeling (Hauser *et al.*, 1989), and differential scanning calorimetry (Hauser *et al.*, 1989; Luisi *et al.*, 1990). Differences with bulk water are more marked at W_0 below 10, while at higher W_0 the properties of micellar water approach those of bulk water. Other NMR studies of micelles formed with phospholipids and cetyltrimethylammonium bromide (CTAB) (Darszon and Shoshani, 1992; Kernen *et al.*, 1997) have provided similar findings. Water molecules closest to the polar head groups of the surfactant molecules exhibit the largest differences from bulk water, due to their strong interactions with the charged polar surface of the micelle.

3B.1.4 Reverse Micelles Exchange Contents

Catalysis occurs when enzyme-filled micelles are mixed with substrate-filled micelles, which suggests that micelles can exchange their contents (Fletcher *et al.*, 1987; reviewed by Bru *et al.*, 1995). Various steps are involved in the transformation of substrate into product: (1) collision of substrate and enzyme-filled micelles, (2) fusion and formation of a transient micellar dimer, (3) exchange of contents, and (4) division of the dimer with regeneration of two micelles. The exchange rate is on the order of 10^6 to 10^8 M^{-1} s^{-1}. It is thought that organic-soluble substrates can partition from the apolar solvent into the water phase of the micelle to be transformed by the entrapped enzyme via a process that is faster than the rate of micellar fusion and division. The rates at which either type of substrate reaches the enzyme are substantially faster than the k_{cat} of most enzyme reactions (1 to 100 s^{-1}), and thus, it can be assumed that for most enzymes, exchange of micellar contents would not be rate limiting for steady-state catalysis (Bru *et al.*, 1995).

3B.1.5 Prospective Studies for Cholesterol Esterase: Probing Intermediate Protein Conformations and Protein-Protein Interactions in Reverse Micelles

In keeping with the overall theme of the present research, namely the use of low water systems to probe protein structure and function, the reverse micellar system can be employed as a tool to inquire into parameters or properties that in biological systems are difficult to determine. Although enzymes in all low water systems exhibit activities that are much lower than in aqueous media in the very low range of water concentrations, the unique features that enzymes exhibit in reverse micelles, and when dispersed in organic solvents (see Chapter 1), provide insight into properties of enzymes that are not easily observed in aqueous solutions.

Reverse micelles can be used to study CE at different levels of association. As mentioned above, the effect of the 3α,7α,12α-trihydroxy bile salts on association of monomer subunits varies with the source of CE [rat pancreatic juice (hexamer); human pancreatic (tetramer); porcine (no effect on association by bile salts but monomer and dimer forms observed); bovine pancreatic (dimer); human milk BAL (no effect on association)]. Monomers of protein at various levels of association have been stabilized only in a few studies (Jaenicke *et al.*, 1981). In this context, various reports with oligomeric enzymes entrapped in reverse micelles look promising and are described briefly below..

Some oligomeric proteins entrapped in reverse micelles exhibit peaks of activity at water concentrations in which the dimensions of the micellar water pool coincide with the size of the subunits that conform the oligomer (Walde *et al.*, 1992; Bru and Walde, 1991, 1993; Mao and Walde, 1991; Bru *et al.*, 1995). For some enzymes, there is conflicting evidence (i.e., maximum activity was observed for some oligomeric enzymes

when the size of the water pool corresponded to the estimated dimensions of the monomer; Kabanov *et al.*, 1991 (see below); Gonnelli and Strambini, 1988). Some enzymes show catalytic rates that are significantly higher than in standard aqueous systems—the so-called superactivity [e.g., α-chymotrypsin (Menger and Yamada, 1979; Barbaric and Luisi, 1981), dihydrofolate reductase (Katiyar *et al.*, 1989), peroxidase (Martinek *et al.*, 1982), wheat germ acid phosphatase (Klyachko *et al.*, 1986), and potato acid phosphatase (Lalitha and Mulimani, 1996)]. It was suggested that in reverse micelles catalytic activity could depend on the activity of micellar water (Gonnelli and Strambini, 1988).

Martinek (1989) and Khmelnitsky *et al.* (1989) have determined the activity of homotetrameric lactate dehydrogenase, entrapped in reverse micelles formed with AOT as a function of Wo. The enzyme activity vs. W_0 profile showed four peaks. The dimensions of the reverse micelles at the four maxima coincided with the dimensions of the enzyme at various states of monomer association. In contrast, the same enzyme entrapped in reverse micelles formed of CTAB (Fernández-Velasco *et al.*, 1992) or AOT (Khmelnitsky *et al.*, 1993a) did not show a peak of maximal activity when examined at various water concentrations. However, the results of time-resolved polarized fluorescence studies of lactate dehydrogenase entrapped in the latter reverse micelles showed that this enzyme could undergo dissociation (Khmelnitsky *et al.*, 1993b).

Activity peaks were also observed with gamma-glutamyl transferase (Kabanov *et al.*, 1989, 1990) and alkaline phosphatase from intestinal mucosa in reverse micelles (Kabanov *et al.*, 1991). The activity peaks appeared in regions in which the size of the water pool of the micelles was equivalent to that of the dissociated monomers and the

data was confirmed by ultracentrifugation studies. The homodimeric enzyme glutathione transferase (octopus) exibited a maximum in plots of activity vs W_o; the size of the water pool at that point corresponded to the estimated dimensions of the monomer (Tang and Chang, 1996). Chemical crosslinking of glutathione transferase entrapped in reverse micelles resulted in no dimers being formed, as evidenced by SDS-PAGE, suggesting that dissociation took place after entrapment.

For α-chymotrypsin and α-chymotrypsin crosslinked to a hexameric structure, a peak of maximal activity was observed in the activity vs W_o profile. However, in the crosslinked enzyme, the peak shifted to a higher W_o, which corresponded to that of micelles with a water pool that had the dimensions of the hexamer (Kabanov *et al.*, 1991).

Moreover, claims that the monomers of oligomeric enzymes entrapped in reverse micelles are catalytically active also appear to be in conflict with data in standard aqueous media, which generally show that the monomers are catalytically inactive. For example, monomers of alkaline phosphatase are catalytically active in reverse micelles (Kabanov *et al.*, 1991; Chang and Shiao, 1994), whereas in water the monomers are inactive (McCracken and Meighen, 1980).

γ-Glutamyl transferase contains a heavy and light subunit (Gardell and Tate, 1981). When the enzyme is entrapped in reverse micelles, it was shown to have dissociated into its constituent subunits from activity data (Kabanov *et al.*, 1989, 1990, 1991). The two monomers were separated by ultracentrifugation and it was observed that the individual monomers could carry out catalysis. This result observed in reverse micelles is especially notable because the active site lies between the two subunits. The

results imply that reverse micelles of surfactants in organic solvents function as matrices with adjustable size permitting to regulate the supramolecular structure and the catalytic activity of oligomeric enzymes. Importantly, reverse micelles offer the possibility for the separation of subunits of oligomeric enzymes under non-denaturing conditions (Kabanov *et al.*, 1991).

Complex reactions such as association of macromolecules in reverse micelles [e.g., restriction enzyme Hind III-phage lambda DNA (Hanley *et al.*, 1991); trypsin-soybean trypsin inhibitor (Bru and García-Carmona, 1991); α-chymotrypsinogen-trypsin (Fadnavis *et al.*, 1993)] also have been explored and some of the observed data may provide information into the molecular events that occur in biological phenomena (Nicot and Waks, 1996).

Research is ongoing to understand the structural arrangements that monomers undergo after association to yield an active oligomer (Green *et al.*, 1995; Thanki *et al.*, 1997). It is worthwhile to explore whether reverse micelles could be a means to determine the structural differences between active and inactive monomers.

Reverse micelles may be used as biological membrane models, and may provide insight into how the structure of CE is affected by its interaction with the surface of an amphiphile, since it physiologically acts at micellar interfaces.

Importantly, intermediates in the formation of native proteins may be stabilized in reverse micelles. Reverse micelles can be used to study the acquisition of the catalytically active three-dimensional structure of CE in "slow motion" (e.g., via time-resolved fluorescence spectroscopy). Bile salts have an obligatory role in the CE activation mechanism. The crystal structures of bovine CE and its complex with

taurocholate have provided a structure-based explanation for the unique role of bile salts in the activation of CE (Wang *et al.*, 1997). There are two bile salt binding sites per CE or BAL monomer. One of these sites is located close to a hairpin loop near the active site. Upon the binding of taurocholate, this loop becomes less mobile and assumes a different conformation. The other bile salt binding site is located remote from the active site. In both structures, BAL forms similar dimers or homodimers with the active sites facing each other. Thus, bile salts activate BAL by binding to a relatively short ten-residue loop near the active site, and stabilize the loop in an open conformation. Presumably, this conformational change leads to the formation of the substrate-binding site, as suggested from kinetic data. The BAL dimer observed in the crystal structure may also play a functional role under physiological conditions. Studying CE entrapped in the low water environment of reverse micelles may provide the unique opportunity to observe the effect of bile salts on the CE intermediates in the formation of the native protein by trapping CE conformers.

On the basis that water is an essential component of conformational mobility, it is theoretically possible to stabilize different protein conformations by adjustment of the amount of water in contact with the enzyme. In Chapter 1, it was shown that enzymes dispersed in organic solvents could form stable conformations that exhibit particular catalytic properties. Studies of enzymes entrapped in reverse micelles have revealed that conformations of enzyme intermediates of biological reactions, which in water have relatively short life times, could be stabilized in the low water environments.

Different water concentrations may be required for an enzyme to undergo different conformational changes to support sufficient flexibility for catalysis. Barrabin

et al. (1993) have shown that different steps of the catalytic cycle of Ca^{2+}-ATPase of the sarcoplasmic reticulum require different water concentrations. The Ca^{2+}-ATPase entrapped in reverse micelles formed with phospholipids catalyzed the hydrolysis of ATP through a process that increased with water concentration. An acyl-phosphate at the catalytic center is an intermediate of the hydrolytic reaction, which can be formed from either phosphate or ATP (De Meis *et al.*, 1980). In comparison to ATP, phosphorylation by phosphate took place with significantly lower amounts of water, indicating that different steps of the catalytic cycle require different water concentrations.

Escamilla *et al.* (1989) have shown that in reverse micelles of the phospholipid type with low water content, the catalytic cycle of cytochrome c oxidase could be arrested at the level of electron transport between hemes a and a_3. Importantly, it is also possible that intermediates of the protein folding pathway could be stabilized in reverse micelles. Lenz *et al.* (1995) observed that insulin labeled with a fluorescent probe entrapped in reverse micelles acquired a molten globule conformation, similar to that observed in solution, which exhibited a weak circular dichroism spectrum in the near UV. reflecting the absence of quaternary structure.

Another line of evidence, which is most relevant with regard to CE, that indicates that intermediates with short-life times can be stabilized in reverse micelles by adjustments of water concentrations derives from experiments that showed that the catalytically active dimer of triosephosphate isomerase was formed after entrapment of unfolded monomers in reverse micelles formed with cetyltrimethylammonium bromide (CTAB) (Garza-Ramos *et al.*, 1992; Fernández-Velasco *et al.*, 1995). The rate at which catalysis appeared in reverse micelles increased with protein concentration, and it was

thus considered that the appearance of catalysis depended on the formation of the dimeric structure. Significantly, similar data in standard aqueous systems have been reported (Waley, 1973; Zabori *et al.*, 1980); however, in reverse micelles at low water concentrations the monomers could be trapped for significantly long times in a state competent for formation of active dimers; in this state a rise in water concentration led to formation of the active dimer. However, data of fluorescence-labeled monomers (Sepulveda-Becerra *et al.*, 1996) did not correlate with the appearance of activity. Experiments with a model system, trypsin and trypsin soybean inhibitor, showed that in reverse micelles with 2% water, protein-protein interactions readily take place (Fernández-Velasco *et al.*, 1995). Fernández-Velasco *et al.* (1995) indicated that at low water concentrations there was a kinetic barrier in the formation of dimers from the individual monomers.

It is apparent from the aforementioned observations that different protein conformations may be trapped by adjustments in water concentration. This may have profound implications on biotechnology and basic protein science, in the sense that an enzyme conformer competent for a given reaction may be stabilized and used for bioconversions, and also because intermediates with a short lifetime in aqueous media may be stabilized and characterized.

3B.1.6 Structure of Enzymes in Reverse Micelles

The structural features of enzymes entrapped in reverse micelles have been reviewed (Nicot and Waks, 1996). Emphasis was made on lipases, myelin basic protein, lysozyme, cytochrome c, liver alcohol dehydrogenase, and serine proteases. Since solutions of reverse micelles with and without proteins are optically transparent,

spectroscopic techniques have been the method of choice for studying protein structure in reverse micelles (Vos *et al.*, 1987; Verhaert *et al.*, 1992); ligand binding has also been employed to determine structural features of proteins entrapped in reverse micelles (Desfosses *et al.*, 1991, 1992).

Circular dichroism (CD) in the far and near UV has often been used to determine the secondary and tertiary structures of proteins entrapped in reverse micelles. Of practical importance was the fact that a problem in the CD measurements, particularly with CTAB, was the strong light absorption of surfactant molecules in the region of 190-240 nm. Verhaert *et al.* (1992) showed that micelles formed with cetyltrimethylammonium chloride (CTAC) instead of CTAB can be used to obtain structural information by far UV CD.

Steady-state fluorescence measurements yield information in the environment of the fluorophore, which can be an intrinsic tryptophan or a covalently bound fluorophore, as evidenced from shifts of fluorescence maxima or changes in quantum yield. Proteins are dynamic entities with fluctuations in times of nanoseconds or seconds. Thus, time-resolved fluorescence provides information of whether the reporter group exhibits restricted motion and undergoes conformational substates. These two techniques have been employed to probe the dynamic structure and environment of the fluorophore (Marzola and Gratton, 1991; Davis *et al.*, 1996; Shoshani *et al.*, 1994; Visser *et al.*, 1994; Shastry and Eftnik, 1996; Dorovska-Taran *et al.*, 1993; Lenz *et al.*, 1995; Strambini and Gonnelli, 1988). Agents that quench the fluorescence of model compounds, or that of intrinsic tryptophan, are employed to probe the characteristics of the region in which the fluorophore is located (Davis *et al.*, 1996).

The secondary structure of α-chymotrypsin entrapped in reverse micelles formed with AOT (W_0 between 6-10) was probed by Fourier transform infrared spectroscopy (FT-IR) (Qinglong *et al.*, 1994). The data showed that entrapment produced a loss of α-helix and β-sheet and an increase in random structure.

On the basis that reverse micelles mimic biological membranes, studies have been conducted on the structural arrangements that some hormones [such as insulin (Lenz *et al.*, 1995)] undergo when they are transferred to reverse micelles. The structure that the hormones acquire in reverse micelles could represent conformations that may exist during biological phenomena (Gallay *et al.*, 1987; Singh and Aruna, 1995).

The overall results suggest that reverse micelles may be a powerful tool to inquire into parameters that in biological systems are difficult to determine.

PART C. Bile Salt-Modulated Stereoselection in the Cholesterol Esterase-Catalyzed Hydrolysis of α-Tocopheryl Acetates

2C. Materials and Methods

Except where stated, all chemicals were purchased from Sigma. *l*-Dimyristoyl-phosphatidylcholine (*l*-DMPC) was purchased from Avanti Polar Lipids Inc. (Birmingham, Ala.). Organic solvents used were HPLC-grade and were purchased from Aldrich or BDH (Toronto, ON).

2C.1 Cholesterol esterases (EC 3.1.1.13; CE)

CE from bovine pancreas (BCE) was purchased from Sigma (#C-1892); protease-free BCE was provided as prepared previously (Moore *et al.*, 1995). Porcine pancreatic CE was purchased from Sigma (#C-9530).

2C.2 Bile Salts

Cholate and taurocholate (sodium salts of cholic and taurocholic acids) were purchased from Sigma; the sodium cholate was recrystallized.

2C.3 Deuterated α-tocopherol derivatives

2*S*,4'*R*,8'*R*-α-(5-CD_3)-tocopheryl acetate (*d_3-SRR*-α-TAc) was previously synthesized (Ingold *et al.*, 1987a, 1987b; Hughes *et al.*, 1990) by the organic synthesis group in the Molecular Selectivity Group of the Steacie Institute for Molecular Sciences, NRC.

2*R*,4'*R*,8'*R*-α-(5,7-(CD_3)$_2$)-tocopheryl acetate (*d_6-RRR*-α-TAc) and

2*R*,4'*R*,8'*R*-α-(5-CD_3)-tocopheryl acetate (*d_3-RRR*-α-TAc) were supplied by Eastman Kodak and by the Natural Source Vitamin E Association.

α-(5,7,8-(CD_3)$_3$)-tocopheryl acetate (*d_9-all-racemic*-α-TAc) and

α-(5,7,8-(CD_3)$_3$)-tocopherol (*d_9-all-racemic*-α-TOH) were previously synthesized by $SnCl_2$-catalyzed deuteriomethylation of α-TAc or α-TOH, for use as an internal standard (Zahalka *et al.*, 1987).

2*R*,4'*R*,8'*R*-α-tocopheryl decanoate (*d_0-RRR*-α-TDec) and

2*R*,4'*R*,8'*R*-α-tocopheryl pentanoate (*d_0-RRR*-α-TPent) were synthesized according to the protocol shown in the synthetic section below. Each compound was used (1) as an internal standard for the GC-MS quantitation of *d_6-RRR*-α-TDec or *d_6-RRR*-α-TPent, respectively, and (2) as a substrate in the HPLC Method for the effect of ester group chain length experiment.

2*R*,4'*R*,8'*R*-α-(5,7-(CD_3)$_2$)-tocopheryl decanoate (*d_6-RRR*-α-TDec) and

2*R*,4'*R*,8'*R*-α-(5,7-(CD_3)$_2$)-tocopheryl pentanoate (*d_6-RRR*-α-TPent) were synthesized according to the protocol shown in the synthetic section below. Each compound was

used as a substrate in the GC-MS method.

PA 322 was a synthetic derivative of vitamin E and was kindly provided for use as an internal standard by Dr. P. Arya of the Steacie Institute for Molecular Sciences of the National Research Council of Canada for the HPLC Method (see below) of analysis for the effect of ester group chain length experiment. This derivative of vitamin E contained a short chain ester function at the "2" chiral position ($CH_2OCOCH_2CH_2CH_2CH_3$) of the phytyl tail of -tocopheryl acetate.

2C.4 Equipment

Spectrophotometric determinations were performed at 302 nm on a Cary 219 UV/visible spectrophotometer with the cell temperature maintained at 37.0 +/- 0.2°C.

α-TOH, α-TAc, and α-TDec were measured on a Hewlett Packard Ultra 1 (crosslinked methyl silicone gum) column (12 m x 0.2 mm x 0.33 μm thickness) in a Hewlett Packard 5890 Gas Chromatograph (280°C, 10 psi for α-TOH- and α-TAc-containing samples, and 20 psi for α-TDec-containing samples) equipped with a 5970 Series Mass Selective Detector. Peaks were detected with UV and MS detectors, which were interfaced, together with the GC-MS autosampler, with a Hewlett Packard 59970C MS Chem Station.

α-TOH, α-TAc, α-TDec, and α-TPent were measured on a Hewlett Packard Series II 1090 Liquid Chromatograph (HPLC) using a hypersil silica column (5 μm in particle size and 200 x 2.1 mm). Peaks were detected with a UV diode array detector. The HPLC, autosampler, and UV detector were interfaced with a DOS Chem Station and data-handling system.

2C.5 Cholesterol Esterase Activity Assay: Initial Rate Kinetics

Exactly 1.5 mL of a micellar *d_3-SRR*-α-TAc solution were placed in an ultraviolet cuvette and incubated at 37°C in a Varian Cary 219 spectrophotometer equipped with a thermostatically controlled cell compartment. A baseline level of α-T-Ac absorbance was measured prior to the addition of an aliquot of cholesterol esterase solution. A few microlitres of enzyme solution were added and the cuvette was shaken briefly. The change in absorbance at 302 nm was monitored (isobestic point, 280 nm). Initial rates of hydrolysis were calculated using the measured difference in molar extinction coefficients at 302 nm between α-T and α-TAc: $\varepsilon_{\alpha\text{-T}} - \varepsilon_{\alpha\text{-TAc}} = 2300\ M^{-1}cm^{-1}$.

Each enzyme preparation (*c*PCE or *pf*BCE) was standardized using the above procedure. A typical enzyme activity assay solution containing 40 mM sodium cholate, 2.0 mM *l*-DMPC in 50 mM Tris buffer/0.15 M NaCl (pH 8) and 0.100 mM *d_3-SRR*-α-TAc was standardized with an aliquot of enzyme which gave an initial rate of hydrolysis of (5.0 +/- 0.1) x 10^{-8} $Msec^{-1}$ at 37°C as monitored at 302 nm. In this way, the enzyme activity could be kept constant throughout all experiments even though the initial enzyme concentration was unknown. For reaction samples of 3.0 mL, the enzyme aliquot was two times that required for the enzyme activity assay, since only 1.5 mL were used in the latter case.

2C.6 Preparation of Micellar Tocopheryl Acetate, Pentanoate or Decanoate Solutions

Individual stock solutions of *d_3-SRR*-α-TAc (1.93 mM), *d_6-RRR*-α-TAc (1.91 mM), *d_0-RRR*-α-TPent (6.55 mM and 1.91 mM), *d_6-RRR*-α-TPent (1.94 mM), *d_0-RRR*-α-TDec (6.298 mM) and *d_6-RRR*-α-TDec (1.91 mM) were prepared in heptane. Stock solutions of 32.0 mM *l*-DMPC were prepared in CH_2Cl_2. Mixtures of:

d_3-SRR-α-TAc + *l*-DMPC, *d_6-RRR*-α-T + *l*-DMPC, *d_3-SRR*-α-TAc + *d_6-RRR*-α-TAc + *l*-DMPC, *d_6-RRR*-α-TDec + *l*-DMPC, *d_0-RRR*-α-TDec + *l*-DMPC, *d_0-RRR*-α-TPent + *l*-DMPC, *d_3-RRR*-α-TAc + *d_6-RRR*-α-TDec + *l*-DMPC were prepared from aliquots of the appropriate stock solutions. The solutions were evapourated under a stream of nitrogen gas and were stored at -20°C until required (up to two weeks). Each sample contained 0.40 μmol (or 100 μM) of *SRR*- and/or 0.40 μmol (or 100 μM) *RRR*-α-tocopheryl acetate, pentanoate or decanoate, and 8.0 μmol (or 2.0 mM) of *l*-DMPC in a final volume of 4.0 mL per reaction sample.

An aliquot of bile salt solution (1.6 mL of 0.1 M sodium cholate or taurocholate) was added to the dry tocopheryl ester/*l*-DMPC mixture and the sample was vortex mixed for 1 min. Aliquots of Tris buffer (0.40 mL of 0.5 M Tris/1.5 M NaCl, pH 8) and distilled deionized water (2.0 mL) were then added, and the solution was vortex mixed after each addition. Samples were then sonicated for 30 min at room temperature to create vesicles/liposomes and ensure the stability of the absorbance at 302 nm. Aliquots (3.0 mL) of the sonicated samples were equilibrated for at least 15 min at 37°C prior to the addition of enzyme (x μL of 1 mg/mL protein cPCE or x μL of *pf*BCE solution which gave a V_{std} = (5.0 +/- 0.1) x 10^{-8} $Msec^{-1}$).

The final concentrations of each component of the aqueous bile salt/buffer solutions with tocopheryl ester were: 100 μM tocopheryl ester in non-competitive experiments, 200 μM tocopheryl ester #1 (100 μM) + tocopheryl ester #2 (100 μM) in competitive experiments, 2 mM *l*-DMPC, 50 mM Tris buffer/0.15 M NaCl, 40 mM bile salt (sodium cholate or taurocholate) (except in the effect of bile salt concentration experiments).

2C.7 Competitive and Non-Competitive Hydrolysis Using Bile Salts

Aliquots (3.0 mL) of *alpha*-tocopheryl ester micellar solutions, prepared as described above, were dispersed into a set of test tubes attached to a manifold (custom glass) maintained under an argon atmosphere. After thermal equilibration of the tocopheryl ester/*l*-DMPC mixed micellar solutions at 37°C (15 min), aliquots of cholesterol esterase were added to each test tube. The aliquot used was the same as that necessary to obtain an initial velocity of approximately (5.0 +/- 0.1) x 10^{-8} $Msec^{-1}$ for the hydrolysis of d_3-*SRR*-α-TAc in sodium cholate/*l*-DMPC mixed micelles (see Section 2C.5). Samples (100 μL) were removed just prior to enzyme addition (t = 0), and at approximately 5, 10, 15, 20, 30, 40, 60, 90, and 120 min for reaction media containing cholate; and at 10, 20, 40, 60, 90, 120, 150, 180, and 210 min for reaction samples containing taurocholate.

Withdrawn samples (100 μL) were placed into test tubes containing 0.5 M SDS (200 μL) to stop the hydrolysis, and a pinch of sodium ascorbate (a reducing agent) to prevent oxidation of the product, α-TOH. Distilled, deionized water (200 μL) was then added to the sample tube which was vortex mixed briefly. The appropriate internal standard solution was then added (i.e. d_9-α-TAc/d_9-α-TOH for experiments on the effect of bile salt concentration, and *l*-DMPC concentration on the rate of hydrolysis of α-TAc stereoisomers; or d_0-*RRR*-α-TDec/d_9-α-TOH for α-TDec hydrolysis reactions; or PA 322 vitamin E synthetic derivative for the comparison of the hydrolysis of Ac/Pent/Dec esters of vitamin E using the HPLC method), and the solution was vortex mixed.

Extraction of the starting material (i.e., α-TAc, α-TPent, or α-TDec), of the product (α-TOH), and of internal standards was then carried out. Ethanol (500 μL) was then added to the sample tube and vortex mixed. Then, 500 μL of heptane, in which the

α-T derivatives are soluble, was added, and the solution was vortex mixed for at least 30 sec to ensure a good extraction. The heptane layer was separated from the aqueous phase using a transfer pipet, and was placed in a small glass vial. The extracted samples were stored overnight at -20°C.

On the following day, the samples were prepared for GC-MS or HPLC analysis. The heptane was evapourated from the samples under a stream of nitrogen gas. The residues were dissolved in 100 μL aliquots of *n*-heptane and vortex mixed. Samples were then transferred into GC-MS plastic vials or HPLC glass vials and sealed with caps equipped with Teflon septa in preparation for GC-MS or HPLC analysis.

2C.8 A. Effect of Bile Salt Concentration on the Rate of *c*PCE Hydrolysis of *SRR*- and *RRR-alpha*-Tocopheryl Acetate Stereoisomers Using *l*-DMPC as the Co-Solubilizing Lipid. Cholesterol Esterase Chiral Selectivity of *SRR*- and *RRR-alpha*-Tocopheryl Acetates Using the GC-MS Method.

(i) Non-competitive (NC) hydrolysis reactions were carried out with either d_3-*SRR*-α-TAc or d_6-*RRR*-α-TAc in the presence of sodium cholate or taurocholate (0, 2, 3, 5, 10, 20, 40 and 80 mM) and *l*-DMPC (2 mM) in Tris buffer (50 mM Tris/0.15 M NaCl), monitoring the absorbance of α-TOH formation from the CEase hydrolysis of α-TAc at 280 nm (GC UV detector) to determine initial velocities of hydrolysis.

(ii) Competitive (C) hydrolysis reactions were carried out on mixtures of d_3-*SRR*-α-TAc and d_6-*RRR*-α-TAc in the presence of 0, 3, 5, 10, 20, 40 and 80 mM sodium cholate or taurocholate and 2 mM *l*-DMPC in Tris buffer (50 mM Tris/0.15 M NaCl).

Aliquots of 1 mg/mL *c*PCE (approximately standardized; see Section 2C.5; range 60-145 μL) were added after the sample incubation period at 37°C. Samples in cholate or taurocholate were then collected at the times mentioned above.

2C.9 B. Effect of *l*-DMPC Concentration on the Rate of *c*PCE Hydrolysis of *SRR*- and *RRR-alpha*-Tocopheryl Acetate Stereoisomers in Cholate or Taurocholate/*l*-DMPC Mixed Micelles Using the GC-MS Method.

Competitive hydrolysis reactions were carried out on mixtures of d_3-*SRR*-α-TAc or d_6-*RRR*-α-TAc in the presence of varying *l*-DMPC concentrations: (i) 2, 4, 6, 8, 10 mM in 5 mM cholate; (ii) 2, 4, 5, 6, 7, 8, 9, 10 mM in 5 mM taurocholate; (iii) 2, 5, 10, 15, 20, 25, 30 mM in 40 mM cholate; (iv) 2, 5, 8, 10, 12, 15 mM in 40 mM taurocholate. All reaction samples contained 50 mM Tris buffer/0.15 M NaCl.

Aliquots of 1 mg/mL *c*PCE (approximately standardized; see Section 2C.5; range 75-100 μL) were added to sample reaction vessels after incubation at 37°C. Samples in cholate or taurocholate were collected at the appropriate times (see above).

2C.10 C. Effect of *alpha*-Tocopheryl Ester Group Chain Length on the Rate of *pf*BCE Hydrolysis of Synthetic *alpha*-Tocopheryl Esters Using the GC-MS Method and the HPLC Method.

(i) Comparison Between d_3-*RRR*-α-TAc and d_6-*RRR*-α-TAc Hydrolysis by *pf*BCE Using the GC-MS Method.

Competitive and non-competitive hydrolysis reactions were carried out with d_3-*RRR*-α-TAc and/or d_6-*RRR*-α-TAc in the presence of 40 mM cholate or taurocholate, 2 mM *l*-DMPC and 50 mM Tris buffer/0.15 M NaCl monitoring absorbance at 302 nm using the GC-MS Method for the determination of initial velocities.

(ii) Comparison Between d_3-*RRR*-α-TAc and d_6-*RRR*-α-TDec Hydrolysis by *pf*BCE Using the GC-MS Method.

Competitive and non-competitive hydrolysis reactions were carried out with d_3-*RRR*-α-TAc and/or d_6-*RRR*-α-TDec esters in the presence of 40 mM cholate or taurocholate, 2 mM *l*-DMPC and 50 mM Tris buffer/0.15 M NaCl.

Aliquots of *pf*BCE (54 μL for samples in cholate and for samples in taurocholate) were added after the samples were incubated at 37°C. Samples in cholate or taurocholate were collected at the appropriate times (see above).

(iii) Comparison Between d_3-*RRR*-α-TAc, d_0-*RRR*-α-TPent, and d_0-*RRR*-α-TDec Hydrolysis by *pf*BCE Using the HPLC Method of Analysis.

Non-competitive hydrolysis reactions were carried out with either d_3-*RRR*-α-TAc, d_0-*RRR*-α-TPent, or d_0-*RRR*-α-TDec synthetic esters in the presence of 40 mM cholate or taurocholate, 2 mM *l*-DMPC and 50 mM Tris buffer/0.15 M NaCl.

Aliquots of *pf*BCE (approximately standardized; see Section 2C.5: 78 μL were used for both samples in cholate and in taurocholate) were added after thermal equilibration of the samples at 37°C. Samples in cholate or taurocholate were taken at the afore-mentioned times, but also at 20 hrs 25 min (the next day). Sample analysis was via the HPLC Method.

2C.11 Analysis of *alpha*-Tocopherol and *alpha*-Tocopheryl Esters

2C.11.1 GC-MS Method

Prepared sample vials were placed in an autosampler (HP 7673 A) connected to a Hewlett Packard 5890 model gas chromatograph (GC). Samples (1 μL) were injected automatically into a Hewlett Packard Ultra 1 fused silica capillary column (12 m x 0.2mm ID; cross-linked methyl silicone bonded phase) maintained at 280°C, and at 10 psi for acetates or phenols of α-T or 20 psi for the decanoate ester of α-T; and eluted with helium gas using a 30:1 split ratio (flow rate). The GC was connected to a Hewlett Packard 5970 A Series Mass Selective (MS) Detector, programmed to monitor continuously the d_0, d_3, d_6 and d_9 molecular ions corresponding to the parent ion masses

of the respective α-T esters or α-TOHs (430.33, 433.33, 436.33, 439.33 atomic mass units for the respective non-deuterated and deuterated α-tocopherols; 472.33, 475.33, 478.33, 481.33 atomic mass units for the respective non-deuterated and deuterated α-tocopheryl acetates). Since the α-TPent and α-TDec esters cannot withstand the high temperature of the GC, they are cleaved at the ester group prior to entry into the MS Detector. Thus, the d_0 and d_6 molecular ions of the parent ion masses of the synthetic α-T esters are at 430.33 and 436.33 atomic mass units, respectively. However, because these long-chained synthetic ester derivatives of α-T are more non-polar than their acetate counterparts, and because the GC method employs a reverse-phase GC column (i.e., where the stationary phase is non-polar as opposed to polar), they will reach equilibration between the stationary phase and the mobile phase (helium gas) relatively more slowly than their acetate and tocopherol analogues. Thus, even though these esters (pentanoate and decanoate) of α-T degrade at high temperatures, their cleaved molecular ions will be detected at a longer retention time than their acetate counterparts that are relatively stable in the gas state. The order of elution using the GC-MS method is: α-T, α-TAc, α-TPent, α-TDec.

The GC, autosampler and MS detector were interfaced with a Hewlett Packard 59970C MS Chem Station, which controlled the analysis and the acquisition of the data. The GC-MS data, in the form of ASCII files, were transferred to an IBM PC clone (386 model) and were analyzed using the Lotus 1-2-3 spreadsheet program.

The concentrations of d_0-, d_3- and d_6-α-tocopheryl acetates, d_0- and d_6-α-tocopheryl decanoates, and their respective tocopherols in the samples were determined by comparing their respective peak areas with the respective peak areas of the

appropriate internal standard, d_9-α-TAc, d_0-*RRR*-α-TDec, or d_9-α-TOH, which were added in known amounts to the appropriate sample just prior to the extraction procedure. All measurements of deuterated α-T derivatives were corrected for the actual degree of incorporation of the nominal amount of deuterium in the molecule. The peak area data for the internal standards were also corrected for the M+3 ion of d_6-α-TAc, OH or Dec. The M+3 contribution is due to the presence of natural abundance isotopes (^{2}H, ^{13}C).

The DOS computer files from the PC were transferred to a MacIntosh computer for graphical analysis and presentation of results using Kaleida Graph, Microsoft Word, and Microsoft Excel programs.

2C.11.2.1 HPLC Method

This method was employed to analyze the samples containing the synthetic esters of α-T in order to speed up the analysis, since the retention time for α-TDec is about 20 min and the total analysis time per sample is 30 min using the GC-MS method. The HPLC method allowed for a total analysis time of approximately 10 min.

Samples were prepared as described above in glass vials, and were placed inside an HPLC autosampler. About 12.5 μL samples were automatically injected into a high performance liquid chromatograph (HPLC; Series II 1090 Liquid Chromatograph) equipped with a hypersil silica normal phase column (polar stationary phase: 5 μm particle size; 200 x 2.1 mm) and eluted with a mobile phase comprised of 99.5% hexane and 0.5% 1,4-dioxane (vol/vol) at a flow rate of 0.5 mL/min. Peaks were detected with a Varian UV diode array detector at 280 nm. The HPLC, autosampler, and detector were interfaced with a DOS Chem Station and data-handling system.

Quantitative analyses were based on the direct relationship between the absolute

amount of solute and the area under the corresponding peak. To relate peak areas to concentrations, an internal standard of known concentration was used. The linearity of the area-concentration (area ratio-amount ratio) relationship was ensured in the concentration range of interest by creating a *calibration curve*. The internal standard method circumvents the problem of irreproducibility of injection volumes.
The internal standard employed was the α-T derivative, PA322, which was eluted in between the retention times of the compounds of interest.

2C.11.2.2 Formation of the calibration curve

Solutions of d_3-*RRR*-α-TAc, d_0-*RRR*-α-TDec, d_0-*RRR*-α-TPent, d_0-*RRR*-α-TOH and PA 322 (50 μM; value kept constant throughout a series of concentrations) were prepared at five different concentrations: 19.1 μM, 38.2 μM, 57.3 μM, 85.95 μM, and 114.6 μM (samples contained 1.0 mL in total volume in each case) for the former three substrates, and 19.2 μM, 38.4 μM, 57.8 μM, 86.4 μM and 115.2 μM for d_0-*RRR*-α-TOH. Five calibration vials were made, each containing approximately the same concentration of each component.

After analysis of the samples as described above, the HPLC DOS computer was used to create calibration curves which related peak area ratios (of the internal standard to each component; i.e., five separate curves were generated, relating the internal standard area ratio to the internal standard amount ratio; correlation factors were in the range R=0.996-1.00) to the amount ratios (of the internal standard to each component). Thus, samples containing unknown concentrations of the components of interest could be analyzed and their peak areas compared with that of the internal standard to form a peak area ratio that is then correlated to an amount ratio from which the analyte concentration

could be obtained.

Since normal phase HPLC was used (i.e., the column stationary phase was polar and the eluent non-polar), the order of elution of the compounds tested was: d_0-*RRR*-α-TDec, d_0-*RRR*-α-TPent, d_3-*RRR*-α-TAc, PA322, d_0-*RRR*-α-TOH, in that the most polar compound was retained the longest on the column, and the least polar compound was eluted first.

Purity checks of each component revealed the following: d_0-*RRR*-α-TOH—97.42%; d_3-*RRR*-α-TAc—98.96%: d_0-*RRR*-α-TPent—100.0 %: d_0-*RRR*-α-TDec—97.1%: PA 322—97.03%.

2C.13 Analysis of Results

Initial rates of reaction were first-order in nature (Fersht, 1984). A first-order reaction occurs when the initial rate of use of substrate (or the rate of product formation) is directly proportional to the substrate concentration:

$$S \xrightarrow{k_f} P$$

$$v = -dS/dt \quad (1)$$

If v is proportional to [S], then $v = k[S]$, (2)

where k is the first-order rate constant.

Therefore, $v = -dS/dt = k[S]$ (3)

Integrating equation (3) gives:

$$-\int dS/S = \int k dt \quad (4)$$

$$\ln[S] = -kt + C \quad (5)$$

where C = a constant.

When $t = 0$, $[S] = [S]_o$ and equation (5) becomes:

$$\ln[S]_o = -k(0) + C$$

Therefore, $C = \ln[S]_o$ and

$$\ln[S]_t = -kt + \ln[S]_o \quad (6a)$$

$$\ln[S]_t/[S]_o = -kt \quad (6b)$$

Since $[S]_o = [P]_\infty$ and (7)

$$[S]_t = [S]_o - [P]_t \text{ then} \quad (8)$$

$$[S]_t = [P]_\infty - [P]_t \quad (9)$$

Therefore, equation (6a) becomes:

$$\ln([P]_\infty - [P]_t) = \ln[P]_\infty - kt \quad (10)$$

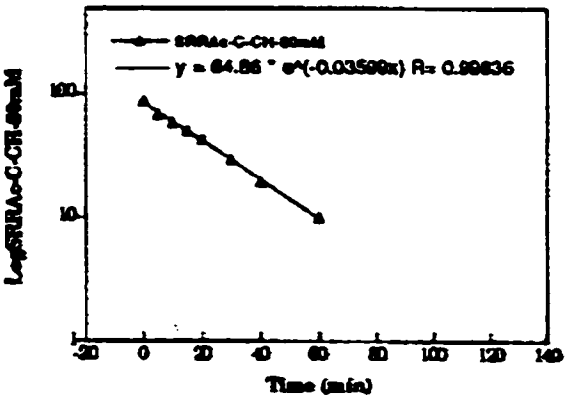
which resembles the equation for a straight line,

$$y = mx + b.$$

Since it was more accurate to determine the initial substrate concentration, $[S]_o$, than to extrapolate the endpoint, $[P]_\infty$, the following equation was used to calculate the first-order rate constants of the hydrolysis reactions (equation (8) was substituted into equation (6a)):

$$\ln([S]_o - [P]_t) = \ln[S]_o - kt \quad (11)$$

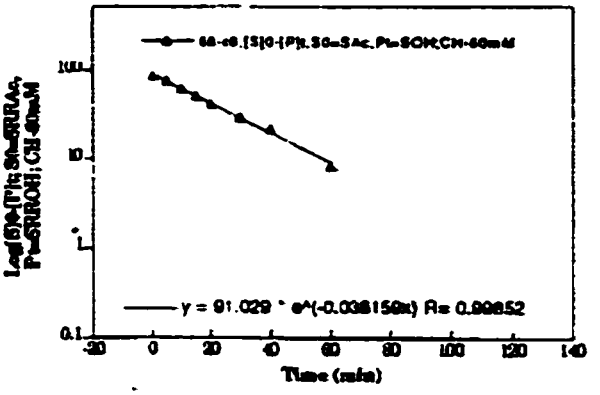
A typical plot of $\text{Log}[S]_t$ vs time looked like the following:



The y-intercept was extrapolated to give $\log[S]_o$. A least squares analysis using an exponential curve fit was used to fit the data directly to the theoretical equation:

$$[S]_t = [S]_o * \exp(-kt) \qquad (12)$$

The extrapolated value for $[S]_o$ was then used in equation (11), and semilogarithmic plots were constructed. A typical plot of $\text{Log}([S]_o - [P]_t)$ vs time looked like the following:



Thus, the [S]o and the first-order rate constant, k, were able to be defined using an exponential curve fit using the following equation:

$$[S]_o - [P]t = [S]_o * \exp(-kt) \quad (13)$$

which rearranges to:

$$[P]_t = [S]_o * (1 - \exp(-kt)) \qquad (14)$$

where P = α-TOH, and S = α-TAc/Pent/Dec.

*At least the first seven points in the product formation curves (up to 80% conversion) could be fitted by semilogarithmic analysis with high correlation coefficients ($R > 0.98$).

2C.14 Synthesis of Substrates

2C.14.1 Step 1: Formation of α-TOH from α-TAc

Due to the fact that α-TOH is oxidized to its quinone form in air, it was necessary to perform a base hydrolysis (saponification; March, 1985a) reaction of α-TAc in order to obtain α-TOH, which was required for subsequent steps toward the synthesis of pentanoate and decanoate esters of vitamin E. *d_6-RRR*-α-TOH was obtained from *d_6-RRR*-α-TAc via the following protocol.

About 10 g of *d_6-RRR*-α-TAc (a yellow, viscous oil) were dissolved in 150 mL of

MeOH. The mixture was stirred until the α-TAc was fully dissolved. K_2CO_3 (5 g) in 10 mL of H_2O were added to the reaction vessel, which was then flushed with N_2 gas to expel O_2 from the vessel. In solution, the K_2CO_3 dissociates to $2K^+ + CO_3^{2-}$. The latter anion attracks the protons from H_2O, and the hydroxide anion thus formed can attack the electrophilic carbon atom of the ester group of *d_6-RRR*-α-TAc. The base hydrolysis or saponification reaction of α-TAc to α-TOH produced fizzing white bubbles. After some time, the solution equilibrated to a yellow, milk-like solution of the potassium salts of *d_6-RRR*-α-TOH and acetic acid.

The reaction was followed by thin-layer chromatography (TLC). An initial TLC developed in a 1:10 mixture of ethyl acetate:hexane gave an R_f(α-TAc)=0.46 and R_f(α-TOH)=0.36. After 3.5 h, a second TLC revealed an R_f(α-TAc) =0.52 and R_f(α-TOH)=0.39 in the same solvent system. The reaction had gone to completion and the vessel was stored in the freezer at -20°C overnight. On the next day, the reaction mixture was worked-up via an acidification procedure. A few millilitres of 1.0 N HCl were added to the mixture until the pH was lowered to neutrality to yield acetic acid and the product, *d_6-RRR*-α-TOH.

The K_2CO_3 was suction filtered off with MeOH. The solvent (MeOH) was evaporated off at room temperature (since MeOH has a b.p. of about 65°C, it can be evaporated off under vacuum without heat), and the viscous oily product, d_6-*RRR*-α-TOH that remained was orangey-reddish-brown. α-TOH was dried using toluene (b.p. 110°C), which forms an azeotrope with water and the latter can be evapourated along with the toluene under vacuum. The sample product was flushed with N_2 gas and stored in the freezer at -20°C overnight. Samples of the *d_6-RRR*-α-TOH product and *d_6-RRR*-α-TAc

were analyzed for their purity using the GC-MS Method at 150°C. The retention time of d_6-*RRR*-α-TOH was approximately 14.4 min, and that of d_6-*RRR*-α-TAc was approximately 15.2 min. The mass spectrum for d_6-*RRR*-α-TOH revealed the molecular ion peak at 437 a.m.u., and the base peak showing the most stable radical cation at 172 a.m.u.; the product was relatively pure, as shown by the GC profile. The mass spectrum for d_6-*RRR*-α-TAc revealed the major base peak at 172 a.m.u., and also a less abundant, but still distinct, minor base peak at 437 a.m.u. The molecular ion peak was very small at 478.95 a.m.u. Thus, the acetate form of α-TOH becomes ionized in the ionization chamber of the mass spectrometer at the ester functionality. The coupling reaction of d_6-*RRR*-α-TOH and decanoic acid was performed on the following day.

2C.14.2 Step 2a: Formation of d_0- and d_6-*RRR*-α-Tocopheryl Decanoate Esters

The synthesis of d_0- or d_6-*RRR*-α-TDec was carried out by the DCC dicyclohexylcarbodiimide (DCC) coupling method (March, 1985b) via the following protocol. One equivalent of DCC can convert one equivalent of decanoic acid and one equivalent of α-TOH to the desired decanoate ester (α-TDec). An equivalent amount of dimethylaminopyridine (DMAP; F.W.122.17; m.p. = 108-110°C) is employed as a catalyst.

From a mechanistic standpoint, the decanoic acid forms an *active ester* with the DCC by nucleophilic attack (by the nucleophilic carboxylic acid oxygen atom of decanoic acid) of the electrophilic carbon atom of DCC. DMAP base catalyst is required to polarize the R'O-H bond so that $\alpha\text{-TO}^{\delta-}\text{-H}^{\delta+}$ becomes a better nucleophile to attack the carbonyl carbon atom of the active ester. The α-TOH attacks the carbonyl carbon atom of the active ester (*O*-acylisourea), in the presence of DMAP, to form a tetrahedral

intermediate that, ultimately, is converted to the product ester (α-TDec) and DCU (dicyclohexyl urea).

A detailed account of the procedure is as follows. To a round-bottomed (100 mL) flask were added 0.17440 g (0.001 mol) of decanoic acid (F.W. 172.27) (b.p.=268-270°C; m.p.=31-32°C; d=0.893 gcm^{-3}), dissolved in about 30 mL of freshly distilled (dry) CH_2Cl_2. The reaction vessel was flushed with N_2 gas. To the above solution were added 0.22734 g (0.0011 mol) of DCC (1,3-dicyclohexyldiimide, F.W. 206.33). The reaction was stirred for 20-25 min or until a cloudy/turbid solution was observed which indicated that the active ester was formed, since it is insoluble in CH_2Cl_2. At this point, it is very important to follow the reaction by TLC (silica gel) to make sure that the *active ester* is formed. Otherwise, when the alcohol (phenol; α-TOH) is added, the coupling reaction will not take place. Also, if a base, R-NH_2, such as DMAP, is added too soon to the reaction mixture, before the decanoic acid has time to react fully with the DCC, a DCC—NH-R unwanted byproduct might form. The same scenario may occur with R'-OH. Thus, the decanoic acid is allowed to react fully with DCC, followed by the addition of α-TOH and DMAP base.

Formation of the active ester (*O*-acylisourea) was monitored by TLC (silica gel), which gave an R_f(active ester of decanoic acid and DCC) = 0.50 in a 1:10 ethyl acetate:hexane solvent system as compared to a spot of the starting solution of decanoic acid and DCC. Decanoic acid does not move up the polar TLC plate. The TLC check showed that there was enough active ester (*O*-acylisourea) formed to proceed with the addition of d_6-*RRR*-α-TOH (0.44787 g or 0.001 mol; F.W.436.67) (dissolved in 5 mL of CH_2Cl_2) and DMAP catalyst. The reaction was stirred overnight.

A TLC of the cloudy reaction mixture before the work-up procedure, developed in 1:10 ethyl acetate:hexane solvent system, revealed a small spot with an R_f(*d_6-RRR*-α-TOH)=0.38, and an R_f(*d_6-RRR*-α-TDec)=0.81; the reaction had gone to completion, and the work-up procedure was carried out. The reaction mixture was washed twice with 60-80 mL of CH_2Cl_2 and with 20 mL of pH 7 buffer solution (1:25 Aldrich pH 7 buffer solution:distilled, deionized H_2O). A separatory funnel was used to extract the CH_2Cl_2 organic layer (bottom layer), which was then dried with $MgSO_4$ and suction filtered. The solvent was evapourated under vacuum, and the crude product was, thus, obtained.

Since an R_f value of about 0.25-0.30 for the product compound is desirable for the purification procedure, it was necessary to find, via a trial and error approach, an appropriate solvent system. A 1:40 ethyl acetate:hexane mixture gave the desired result: R_f(*d_6-RRR*-α-TOH)=0.11 and R_f(*d_6-RRR*-α-TDec)=0.27.

Purification of the product was carried out via flash chromatography with 1:40 ethyl acetate:hexane (Still *et al.*, 1978). The fractions containing the product were collected, the solvent was evapourated under vacuum, and the pure sample was obtained. Crude and pure samples of *d_6-RRR*-α-TDec were analyzed for purity via GC-MS at 280°C. The retention time of the crude sample of *d_6-RRR*-α-TDec was approximately 12.5 min, and that of the pure sample was at about 12.4 min. The mass spectrum of the crude sample showed the base peak at 172 a.m.u. and the second most stable radical cation was evident at 437 a.m.u.. For the pure sample, the mass spectrum revealed the latter two peaks, as well as a very small molecular ion peak at 591 a.m.u. (M.W. of *d_6-RRR*-α-TDec is 590.93).

Exactly the same procedure was used for the synthesis of d_0-*RRR*-α-TDec. The TLC check for the presence of active ester at 80 min after the DCC was added to the decanoic acid showed an R_f(active ester) = 0.25 developed in a 1:10 ethyl acetate:hexane solvent system. A 1:7 ethyl acetate:hexane solvent system was used to obtain a good separation of the product and the starting material. The TLC revealed an R_f(d_0-*RRR*-α-TOH)=0.54 and an R_f(d_0-*RRR*-α-TDec)=0.92. Flash chromatography of the reaction product, d_0-*RRR*-α-TDec, was first carried out with a 1:40 ethyl acetate:hexane solvent system; however, since the separation was not very good, a more non-polar solvent system was required to completely separate the product from the remaining impurities (e.g. starting material, d_0-*RRR*-α-TOH). The 1:60 ethyl acetate:hexane solvent system was chosen; however, prior to the sample application onto the column, the silica gel was equilibrated with pure hexane; and, then, the sample was introduced onto the column, and eluted with the 1:60 ethyl acetate:hexane solvent system. This ensured that the sample would not quickly pass through the column before it had a chance to separate from the impurities. The GC-MS results are as follows: the retention time of d_0-*RRR*-α-TOH was approximately 13.4 min (at 150°C), whereas, that for d_0-*RRR*-α-TDec was about 17.3 min (at 280°C); the mass spectrum of d_0-*RRR*-α-TOH revealed a base peak at 165 a.m.u. and the molecular ion peak at 430 a.m.u., and that of d_0-*RRR*-α-TDec showed a base peak at 164 a.m.u and the second most stable peak at 430 a.m.u. (the molecular ion peak was not detected by the MS Detector since the ester is unstable in the gas state).

2C.14.3 Step 2b: Formation of d_0- and d_6-*RRR*-α-Tocopheryl Pentanoate Esters

The synthesis of d_0- or d_6-*RRR*-α-TPent was carried out by the following procedure. d_0- or d_6-*RRR*-α-TOH (0.002 mol of each; i.e., 0.86303 g of d_0-*RRR*-α-TOH

(M.W. 430.67) or 0.88980 g of d_6-*RRR*-α-TOH (M.W. 436.67)) were dissolved in 35 mL of dry CH_2Cl_2 and the flask was flushed with N_2 gas. Triethylamine (Et_3N) was then added to the α-TOH solution via a syringe (0.003 mol or 0.42 mL Et_3N; M.W. 101.1912; d=0.726 g/mL; m.p.=-115°C; b.p.=88.8°C), which served the purpose of making the α-TOH a better nucleophile by abstracting its proton. Valeryl chloride ($CH_3(CH_2)_3COCl$; F.W. 120.58; b.p.=125-127°C; d=1.016 g/mL), an activated acid chloride, was then added to the reaction vessel (0.0022 mol or 0.27 mL) via a syringe.

The reactions were followed by TLC. Both reactions were run at the same time, and after 2.5 h., the TLCs for d_0-*RRR*-α-TPent and d_6-*RRR*-α-TPent, developed in 1:7 ethyl acetate:hexane, revealed the following R_f values: R_f(d_0-*RRR*-α-TOH)=0.52 and R_f(d_0-*RRR*-α-TPent)=0.79; and R_f(d_6-*RRR*-α-TOH)=0.48 and R_f(d_6-*RRR*-α-TPent)=0.77. The solvent systems chosen for the purification procedure (via flash chromatography) of both d_0-*RRR*-α-TPent and d_6-*RRR*-α-TPent were the same; and each column was first equilibrated with pure hexane. The R_f values in 1:30 ethyl acetate:hexane were: R_f(d_0-*RRR*-α-TOH)=0.15 and R_f(d_0-*RRR*-α-TPent)=0.28; and R_f(d_6-*RRR*-α-TOH)=0.22 and R_f(d_6-*RRR*-α-TPent)=0.34. Purified samples were tested for purity by GC-MS at 280°C. The retention times for d_0-*RRR*-α-TPent and d_6-*RRR*-α-TPent were at 4.2 min and 5.9 min, respectively. This difference in retention times was probably due to slightly different pressure conditions, as the GC-MS samples were tested on different days. The mass spectra revealed the following: for d_0-*RRR*-α-TPent, the base peak was at 165 a.m.u., the second most stable radical cation peaks were at 430 a.m.u. (d_0-*RRR*-α-TPent is degraded in the gas state, before it reaches the MS Detector) and 57 a.m.u., and the very small molecular ion peak was evident at 514 a.m.u.; for

d_6-*RRR*-α-TPent, the base peak was at 435 a.m.u., the second most stable radical cation peaks were at 170 a.m.u. and at 56 a.m.u., and the very small molecular ion peak was at 520 a.m.u..

3C. Results

3C.1 Noncompetitive and Competitive Hydrolyses of d_6-*RRR*-α-TOAc and d_3-*SRR*-α-TOAc Catalyzed by PCE and Trihydroxy Bile Salts

Figure 6-8 shows the kinetic traces of α-TOH formation as a function of time for competitive hydrolyses of d_6-*RRR*-α-TOAc and d_3-*SRR*-α-TOAc by PCE in cholate and taurocholate bile salt/*l*-DMPC mixed micelles in which the bile salt concentration was varied. The non-competitive plots (not shown) had relatively higher initial rates of hydrolysis than in the competitive cases. The direction of chiral discrimination between *RRR*- vs *SRR*-α-TOAc in terms of their relative rates of hydrolysis depends on the bile salt used and on whether the experiment was run in a noncompetitive or competitive manner (Figure 6-9 and Figure 6-10). PCE discriminates in favour of the *SRR*-α-TOAc substrate in cholate/*l*-DMPC mixed micelles, but the *RRR*-α-TOAc substrate is favoured in taurocholate/*l*-DMPC mixed micelles. The difference in selectivities found in the competitive and non-competitive experiments have been attributed to competitive inhibition by diastereomeric phenol products (Zahalka *et al.*, 1991).

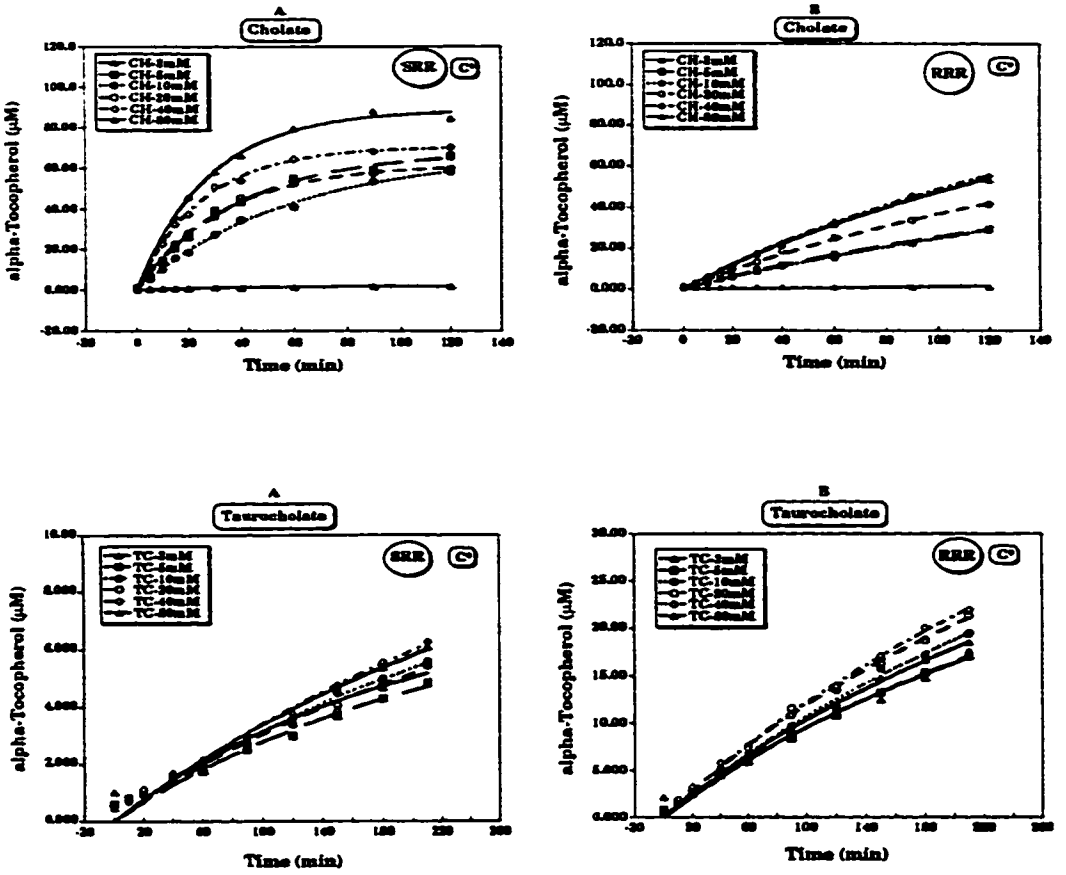


Figure 6-8. Effect of cholate (CH) and taurocholate (TC) concentration on α-TOH product formation during the competitive (C*) hydrolysis of (A) d_3-*SRR*-α-TAc and (B) d_6-*RRR*-α-TAc by cPCE in bile salt/*l*-DMPC mixed micelles. Please note the difference in ordinate scales for the TC cases A and B.

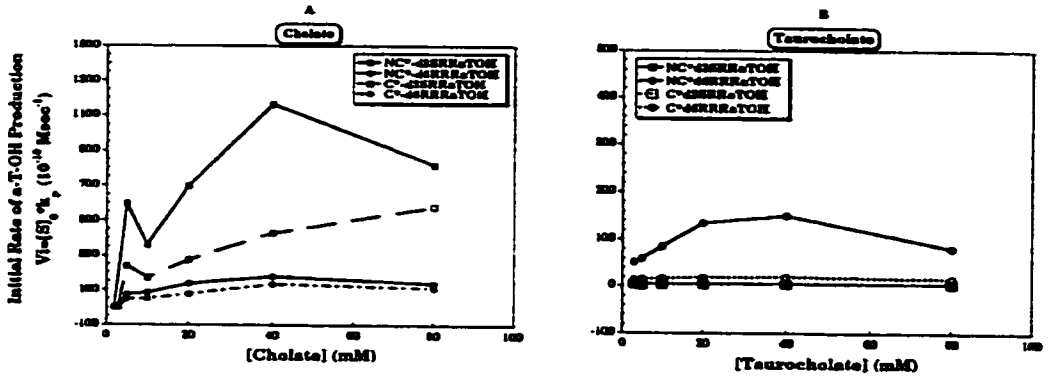


Figure 6-9. Effect of (A) cholate and (B) taurocholate concentration on the initial rates of non-competitive (NC*) and competitive (C*) hydrolyses of d_3-*SRR*-α-TAc and d_6-*RRR*-α-TAc by cPCE in bile salt/*l*-DMPC mixed micelles.

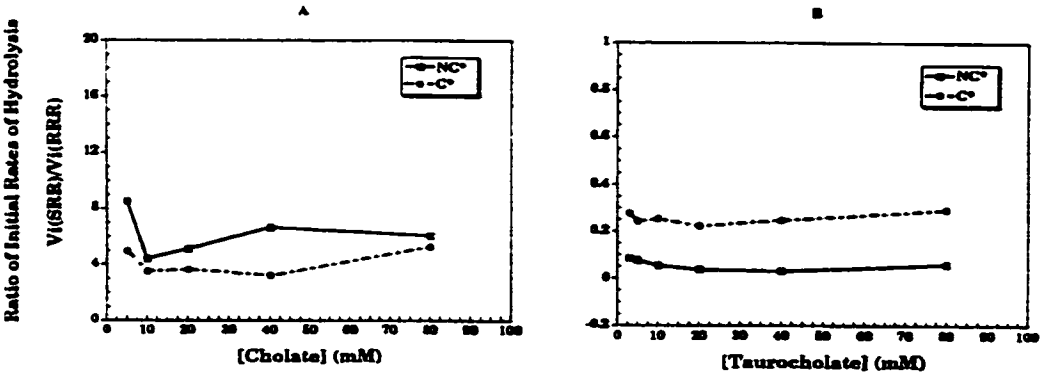


Figure 6-10. Effect of (A) cholate (CH) and (B) taurocholate (TC) concentration on the ratios of the initial rates on non-competitive (NC*) and competitive (C*) hydrolyses of d_3-*SRR*-α-TAc and d_6-*RRR*-α-TAc by cPCE in bile salt/*l*-DMPC mixed micelles.

3C.2 Effect of Bile Salt Concentration

Table 6-2 shows the effect of the concentration of the trihydroxy bile salts cholate and taurocholate on the initial rates for the cPCE-catalyzed noncompetitive and competitive hydrolyses of *RRR*- vs *SRR*-α-TOAc under standard conditions (including 2.0 mM *l*-DMPC). There was no measurable reaction in the complete absence of bile salt, but reaction could be readily detected upon the addition of as little as 3 mM of these CE cofactors. In both the noncompetitive and competitive experiments the reaction rates do not change dramatically over a range of bile salt concentrations from 5 to 80 mM (Figures 6-9 and 6-10). The direction and roughly the magnitude of the chiral discrimination induced by each bile salt are not greatly dependent on the bile salt concentration.

3C.3 Effect of *l*-DMPC Concentration

Competitive cPCE-catalyzed hydrolyses were carried out using 5 mM cholate or 5 mM taurocholate with 2, 4, 6, 8, and 10 mM *l*-DMPC. Experiments with higher bile salt concentration (the normal 40 mM) were also carried out, but there were little differences in the diastereoselectivity toward *RRR*- vs *SRR*-α-TOAc compared to the experiments carried out with 5 mM bile salt. Initial rates are given in Table 6-3. The 5 mM bile salt concentration was chosen for the experiments in order to examine the results more readily at high [*l*-DMPC]/[bile salt] ratios. The limit of *l*-DMPC concentration at which there occurs very little hydrolysis is observed at a lower value in the 5 mM cholate system ([*l*-DMPC] = 10 mM) cf. in the 40 mM cholate system ([*l*-DMPC] = 30 mM). The latter system has a bile salt/*l*-DMPC ratio of 1.3, which reflects a "dilution effect" of the lipid on the bile salt-lipid mixed micellar phase, and

indicates that a different phase structure altogether is present in such conditions. Similarly, in 5 mM taurocholate very little hydrolysis occurred at 10 mM *l*-DMPC; it was difficult to obtain initial velocity values for 7-9 mM *l*-DMPC from semilogarithmic plots because the data was too scattered, which suggests that the limit of *l*-DMPC concentration in 5 mM taurocholate is in the 5-6 mM range (ratio of bile salt/lipid is 1.0). Cloudy solutions were formed when the concentration of *l*-DMPC reaches 6 mM for both bile salt systems; the mixture forms an emulsion rather than a clear micellar solution. Although this causes the rates of hydrolysis to decline somewhat, the degree of diastereoselectivity appears to be unaffected.

3C.4 Effect of α-Tocopheryl Ester Group Chain Length on the Initial Rates of Hydrolysis by Bovine Cholesterol Esterase

Figure 6-11 illustrates the effect of α-T ester group chain length on the initial rate of α-TOH product formation during the noncompetitive hydrolysis of d_3-*RRR*-α-TAc (acetate), d_0-*RRR*-α-TPent (pentanoate) and d_0-*RRR*-α-TDec (decanoate) by BCE in 40 mM cholate or taurocholate bile salt/*l*-DMPC mixed micelles. In both bile salt mixed micellar systems, the order of hydrolysis by BCE from fastest to slowest is Pent>Ac≥Dec (Table 6-4). The product formation curves for d_3-*RRR*-α-TAc yield a normal noncompetitive conversion to α-TOH of 90% in cholate, and 80% in taurocholate bile salt/*l*-DMPC mixed micelles. Those for d_0-*RRR*-α-TPent and d_0-*RRR*-α-TDec yield a noncompetitive conversion to α-TOH of 60% and 30%, respectively, in both cholate and taurocholate mixed micelles.

Table 6-2. Effect of Trihydroxy Bile Salt Concentration on the Initial Rates of the cPCE-Catalyzed Hydrolysis, V_i, of *RRR*-α-TOAc (100 μM) and *SRR*-α-TOAc (100 μM) in Noncompetitive and (Competitive) Experiments[a]

Bile Salt	[Bile Salt] (mM)	$x10^{-8}$ V_i (Ms^{-1})[b] *RRR*	$x10^{-8}$ V_i (Ms^{-1})[b] *SRR*	V_i^{RRR}/V_i^{SRR} [c]
cholate	0	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	3	0.17 (0.25)	0.26 (--)	0.66 (--)
	5	0.81 (0.43)	5.92 (2.49)	0.14 (0.17)
	10	0.79 (0.47)	3.60 (1.67)	0.22 (0.28)
	20	1.41 (0.79)	6.85 (1.77)	0.21 (0.45)
	40	1.87 (0.99)	10.91 (4.12)	0.17 (0.24)
	80	1.38 (1.05)	8.21 (5.09)	0.17 (0.21)
taurocholate	0	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	3	0.62 (0.23)	0.046 (0.083)	13.53 (2.71)
	5	0.78 (0.22)	0.11 (0.15)	7.13 (1.49)
	10	0.97 (0.31)	0.10 (0.16)	9.35 (1.88)
	20	1.60 (0.37)	1.66 (0.0.22)	-- (1.69)
	40	1.57 (0.30)	0.16 (0.15)	9.82 (2.08)
	80	0.96 (0.31)	0.14 (0.18)	6.97 (1.70)

[a] Except for the bile salt concentrations (when not equal to 40 mM), the experimental conditions were the same as those for standard conditions: 37°C, pH 8.0 with 40 mM bile salt, 50 mM Tris buffer, 2 mM *l*-DMPC and 100 μM α-TOAc in the non-competitive experiments and 200 μM α-TOAc ([*RRR*] = [*SRR*] = 100 μM) in the competitive experiments. Aliquot of CE standardized solution (see Section 2C.5) was used.
[b] Rates calculated from semi-logarithmic plots (see Section 2C.13). Individual rate measurements were reproducible to ±5%.
[c] Errors for non-competitive selectivity ratios are ±10% and for competitive selectivity ratios are ±5%.

Table 6-3 Effect of *l*-DMPC Concentration on the Initial Rates of the cPCE-Catalyzed Hydrolysis, V_i, of *RRR*-α-TOAc (100 μM) and *SRR*-α-TOAc (100 μM) in (Competitive) Experiments[a]

Bile Salt	[*l*-DMPC] (mM)	x10^{-8} V_i (Ms^{-1})[b]	x10^{-8} V_i (Ms^{-1})[b]	V_i^{RRR}/V_i^{SRR} [c]
		RRR	***SRR***	
cholate (5 mM)	2	(0.79)	(3.07)	(0.26)
	4	(0.50)	(1.94)	(0.26)
	6	(0.40)	(1.82)	(0.22)
	8	(0.11)	(0.56)	(0.19)
	10	(0.083)	(0.13)	(0.62)
taurocholate (5 mM)	2	(0.15)	(0.09)	(1.71)
	4	(0.20)	(0.062)	(3.21)
	6	(0.083)	(0.092)	(0.93)
	8	(--)	(--)	(--)
	10	(0.39)	(0.021)	(1.86)
cholate (40 mM)	2	(0.99)	(3.55)	(0.28)
	5	(0.91)	(2.63)	(0.34)
	10	(0.64)	(1.79)	(0.36)
	15	(0.26)	(0.73)	(0.36)
	20	(0.17)	(0.63)	(0.28)
	25	(0.074)	(0.35)	(0.21)
	30	(0.37)	(0.44)	(0.84)
taurocholate (40 mM)	2	(0.50)	(0.30)	(1.65)
	5	(0.25)	(0.088)	(2.86)
	8	(0.11)	(0.023)	(4.70)
	10	(0.13)	(0.11)	(1.30)
	12	(0.045)	(0.039)	(1.15)
	15	(0.077)	(0.030)	(2.57)

[a] Except for the bile salt concentration (when not equal to 40 mM) and the *l*-DMPC concentration (when not equal to 2 mM), the experimental conditions were the same as those that generated the competitive data given in Table 6-2. [b] See footnote *b* in Table 6-2. [c] See footnote *c* in Table 6-2.

Table 6-4 Effect of Ester Group Chain Length on the Initial Rates of the *pf*PCE-Catalyzed Hydrolysis, V_i, of *d_3-RRR*-α-TAc, *d_0-RRR*-α-TDec and *d_0-RRR*-α-TPent in Noncompetitive Experiments[a]

Substrate	$x10^{-8}$ V_i (Ms^{-1})[b]
	RRR
Cholate (40 mM)	
d_3-RRR-α-TAc	1.52
d_0-RRR-α-TPent	6.66
d_0-RRR-α-TDec	0.66
Taurocholate (40 mM)	
d_3-RRR-α-TAc	0.40
d_0-RRR-α-TPent	2.79
d_0-RRR-α-TDec	0.21

[a] The experimental conditions were the same as those for standard conditions: 37°C, pH 8.0 with 40 mM bile salt, 50 mM Tris buffer, 2 mM *l*-DMPC and 100 μM α-T ester. Aliquot of CE standardized solution (see Section 2C.5) was used. [b] See footnote *b* in Table 6-2.

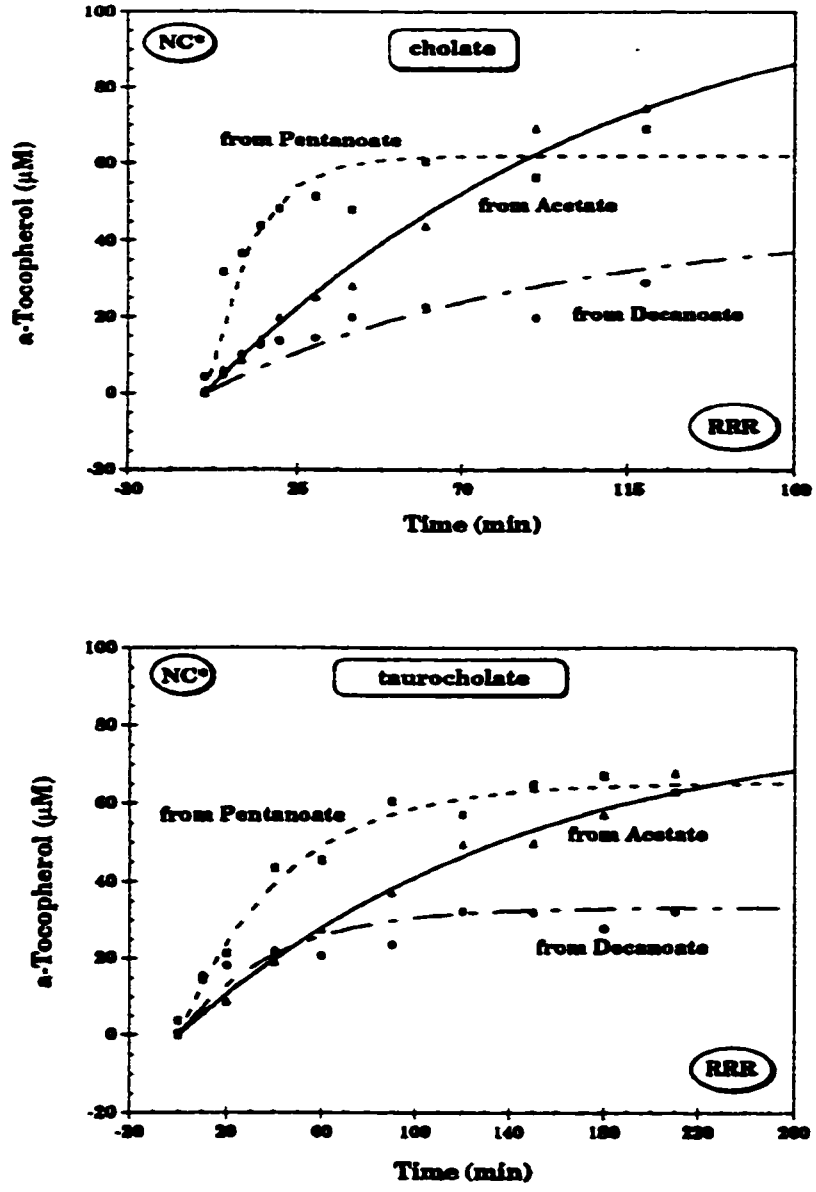


Figure 6-11. Effect of α-T ester group chain length on the α-TOH product formation during the non-competitive (NC*) hydrolyses of *d_3-RRR*-α-TAc, *d_0-RRR*-α-TPent and *d_0-RRR*-α-TDec by *pf*BCE in 40 mM (A) cholate and/or (B) taurocholate bile salt/ *l*-DMPC mixed micelles.

3C. Discussion

The *in vitro* results with cholate and taurocholate confirmed that the diastereoselectivity for the CE-catalyzed hydrolysis of *RRR*- vs *SRR*-α-TOAc could be modulated by the choice of bile salt used to activate the enzyme (Zahalka *et al.*, 1991). Moore *et al.* (1995) have carried out a complete study on this work including glycocholate and dihydroxy bile salts. The discovery of bile salt-modulated diastereoselectivity is unusual and exciting because a major characteristic of catalysis by enzymes is their stereoselectivity (Fersht, 1985). The chiral selectivity of an enzyme can be modified if the precise structure of the active site is altered. This can be achieved by protein engineering (Fersht *et al.*, 1984; Knowles, 1987; Atkins and Sligar, 1989), by unfolding and refolding the enzyme (Wu *et al.*, 1990), or by changing the reaction medium (Westcott and Klibanov, 1994). In the present study, chiral selectivity was altered by the structure of the surface-active compound (i.e., the bile salt) that is used to solubilize the substrate and activate the enzyme.

The two trihydroxy bile salts used in this study activate CE and thus allow the hydrolysis of both α-TOAc diastereomers to occur. The rates of hydrolysis in both the noncompetitive and competitive experiments are much lower when the bile salts are used without a co-lipid than when *l*-DMPC is present together with the bile salt. However, the presence or absence of a co-lipid and its quantity relative to the obligatory bile salt only affects the reaction rates, they do not produce a significant difference in the direction of the diastereoselectivity for *RRR*- vs *SRR*-α-TOAc induced by the bile salt in either the noncompetitive or the competitive experiments. This suggests that the detailed structure of the bile salt micelle does not determine the stereoselectivity. Rather, the bile salts

would appear to modulate the diastereoselectivity of CE by a direct effect on the protein that may involve a "refolding" of the enzyme, with such a "refolding" producing a change in the shape of the active site. Wang and Hartsuck (1993) have suggested that bile salts may cause a conformational change in BAL or CE to provide active site access for the bulky substrate molecule. The actual hydrolysis undoubtedly occurs via a typical serine hydrolase mechanism (Rudd and Brockman, 1984; Lombardo and Guy, 1981; Stout *et al.*, 1985; Sutton *et al.*, 1990; Blow, 1976; Stroud, 1974; Lombardo, 1982; DiPersio *et al*, 1991). Experimental work on human CE (Lombardo and Guy, 1980) and on the bile salt-stimulated human milk lipase (Bläckberg and Hernell, 1993) has led to the conclusion that there are two bile salt binding sites on each of these enzymes. There is some disagreement about the class of bile salts (pattern of hydroxylation) that bind to each site, it seems possible that one site activates the enzyme and controls the stereoselectivity, while the second site encourages the binding of the CE to the surface of micellized lipids. The interfacial binding of CE to the bile salt/lipid mixed micelle is, presumably, mainly or entirely due to the electrostatic attraction between the negatively charged bile salts at the micellar surface and a positively charged CE surface (probably provided by arginine residues) (Lombardo *et al.*, 1983).

The differences between the noncompetitive and competitive diastereo-selectivities are likely due to the fact that competitive inhibition of the enzyme by one diastereomeric acetate or its phenolic product can, in the competitive experiment, retard hydrolysis of the other acetate diastereomer. It is noteworthy that with cholate the PCE-catalyzed hydrolyses of *RRR*- vs *SRR*-α-TOAc occur at very roughly similar rates in the noncompetitive and competitive experiments (Table 6-2). With taurocholate, this is

not true for the PCE-catalyzed hydrolysis. Instead, there is a drastic decline in the rate of the PCE-catalyzed hydrolysis of *RRR*-α-TOAc in the competitive experiment but not in the rate of hydrolysis of *SRR*-α-TOAc. This leads to much lower diastereoselectivities in the competitive experiments than the selectivities calculated from the noncompetitive rate data (Table 6-2). Moore *et al.* (1994, 1995) have shown that glycochenodeoxycholate (a dihydroxyl bile salt containing no –OH group at the 12α position) is very effective in activating PCE (and not BCE). This bile salt/CE combination induces the hydrolyses of the two diastereoisomeric esters at essentially the same rate in the competitive experiment as in the noncompetitive experiments. As a consequence, this bile salt/CE combination produces a uniquely high diastereoselectivity in the competitive experiment. Moore *et al.* (1995) have suggested that the observed diastereoselectivities for hydrolyses in noncompetitive and competitive experiments are controlled partly or entirely by differences in the rates of escape of the diastereomeric phenol products.

The observation that bile salt modulation of the diastereoselectivity of the CE-catalyzed hydrolyses of *RRR*- vs *SRR*-α-TOAc is quite remarkable. Competitive experiments are far more significant than the noncompetitive experiments if the bile salt-modulating effect on CE-catalyzed reactions is to be exploited in organic syntheses. The diastereoselectivity, V_i^{RRR}/V_i^{SRR}, observed in competitive experiments varies from a high of 8.0 for the glycochenodeoxycholate/PCE couple (Moore *et al.*, 1995) to a low of 0.21 for the cholate/PCE couple (Table 6-2). Such a dramatic 40-fold change in the chiral selectivity of an enzyme is without precedent, and as yet explanation.

The chiral carbon atom of *RRR*- or *SRR*-α-TOAc is separated by six bonds from the bond that is broken. The chiral center must exert its effect either prior to, or during,

the rate-limiting step in the overall hydrolysis, probably during the formation of the α-T-ester/CE tetrahedral intermediate. The fact that it is the bile salt alone that determines the direction and magnitude of the *RRR*- vs *SRR*-α-TOAc diastereoselectivity, and not the nature of the co-lipid nor the bile salt/co-lipid ratio, implies that the selectivity is probably not determined by the precise surface structure of the micelle on which the enzyme sits and acts.

The results of the study of the rate of hydrolysis of *RRR*-α-tocopheryl esters of varying ester group chain length by BCE revealed that in both cholate or taurocholate mixed micellar systems the rate of hydrolysis of the *RRR*-α-T esters increased in the order Pent>Ac≥Dec and in each system the extent of *RRR*-α-TOH formation increased in the order Dec<Pent<Ac.

The acetate, pentanoate, and decanoate esters of *RRR*-α-TOH are presented to CE for hydrolysis in mixed micellar form containing bile salts, *l*-DMPC and the appropriate substrate. According to the complex kinetic model for the surface/interfacial reaction of enzymes, one assumption made is that the reaction products must be water-soluble in that they would desorb from the interface, and diffuse rapidly away from the interface; and therefore, do not induce any change with time in the physico-chemical properties of the interface (Ransac, 1990). The *RRR*-α-TOH product is thought to remain in the substrate "superstructure" micellar form after the hydrolysis has occurred because it is amphiphilic in nature (Ransac, 1990). Depending on the size of the acid that is generated, the acid product (i.e., acetic, pentanoic or decanoic acid) either partitions into the aqueous phase (as with short-chained acids such as acetic acid) or remains in the mixed micelle (as is thought to be the case with pentanoic and decanoic acids). Hence, increasing product

concentrations would continuously modify the chemical composition of the interface as catalysis occurs. The water-insoluble amphiphilic product compounds that remain in the mixed micellar interface along with the detergent (the bile salts) and the substrate (*RRR*-α-T esters) may inhibit the hydrolysis reaction (Ransac, 1990). Furthermore, this transient product accumulation in the interface could influence the enzyme velocity as well as other kinetic parameters such as interfacial enzyme concentration, kinetic constants, etc. (Ransac, 1990).

Sutton *et al.* (1990) have studied the effect of varying the chain length of *p*-nitrophenyl ester (PNP esters) substrates on the rate of hydrolysis by CE. Kinetic and nucleophilic trapping experiments (addition of nucleophiles (e.g., CH_3ONH_2, NH_2OH, ethylenediamine and imidazole) activates CE-catalyzed reactions by providing an alternate pathway for forward flux through the acylenzyme intermediate) have shown that k_{cat} is rate-limited by deacylation across the homologous series of substrates. The dependence of k_{cat}/K_m on acyl chain length displays a maximum for the C_6 ester. Both the ascending and descending limbs of the structure-reactivity profile gave linear free energy plots of $\ln(k_{cat}/K_m)$ versus the number of acyl carbons of substrate, with slopes that yield $\Delta\Delta G^{\ddagger} = -430$ (for C_2-C_6 substrates; indicates that the substrate is partitioning from an aqueous to a more apolar environment on conversion of the reactants, i.e., free substrate and micelle-bound free enzyme, to the acylation transition state) and 470 cal/mol per methylene unit (for C_6-C_{12} substrates; indicates a change from a more apolar to a less apolar or more aqueous environment on conversion of the reactants to the transition state), respectively. Solvent isotope effects ($^{D_2O}k_{cat}/K_m$) were shown to decrease from about 2 for short esters to about 1.2 for long esters. It was proposed that

the energetic cost of extraction of the substrate into the CE active site is sufficiently high that substrate binding, a physical step, controls the rate. This process was not expected to generate much of a solvent isotope effect; indeed, the solvent isotope effect was observed to be very small (approximately 1.2) for C_8-C_{10} substrates, and the process is in accord with the descending limb of the free energy plot. Thus, for the long substrates residing primarily in mixed micelles with Triton X-100, the chemical transition state does not contribute appreciably to acylation rate determination. Hence, with increasing chain length, parallel pathways that involve aqueous and micelle-bound pools of substrate operate, and for the longest substrates, a single route that involves rate-determining substrate binding from the micelle pool operates. Except for the C_2 and C_3 esters, proton inventories of k_{cat}/K_m were shown to be linear, which suggests that there is not a change in the rate-determining step for C_4 and longer substrates. These results indicate that the phenomenological transition state for the acylation stage of CE catalysis is highly variable. For short substrates (C_2 and C_3), serial microscopic transition states contribute to rate determination. For the C_4 substrate, a single transition state that is stabilized by a single general-acid-base proton transfer is rate-determining, while with increasing acyl chain length rate determination shifts between parallel reaction pathways. The deacylation rate constant k_{cat} is nearly invariant for C_4-C_8 substrates, but drops off sharply for C_{10} and C_{12} substrates. (The trend in k_{cat} values with increasing fatty acyl chain length defines the structural dependence of the lifetime of acylenzyme intermediates.) However, solvent isotope effects ($^{D_2O}k_{cat}$) are about 2, and proton inventories are linear for all substrates. Therefore, both the acylation and deacylation stages of CE-catalyzed hydrolysis of lipid *p*-nitrophenyl esters have chemical transition states that are stabilized

by single proton transfers (Sutton *et al.*, 1990).

As with the PNP ester substrates, the α-T ester substrate of medium ester group chain length (i.e., α-TPent) was hydrolyzed at a faster rate than α-T ester substrates of short (i.e., α-TAc) or long (i.e., α-TDec) ester group chain length. Even though lipid PNP esters are somewhat distant structurally from the physiological substrates of CE, the deacylation stage of the catalytic mechanism has been shown to be identical with that of physiological CE catalysis (Sutton *et al.*, 1990). Similar k_{cat} values for *p*-nitrophenyl decanoate and for the *sn*-1 thioester of didecanoyl-phosphatidylcholine suggest that deacylation is rate-determining for both, since both substrates produce the same decanoyl-CE intermediate (Sutton *et al.*, 1990). Hence, the work of Sutton *et al.* (1990) supports the observations of the present study.

4. Acknowledgements

The work on the purification of CE and bile salt-modulated stereoselection of α-tocopheryl acetates was supported in part by the National Foundation for Cancer Research. The work on the characterization of taurocholate-induced dimerization of BCE was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

5. References

5A. References

PART A. Purification of Bile Salt-Activated Lipases

Abouakil, N., Rogalska, E., Bonicel, J., and Lombardo, D. (1988). *Biochim. Biophys. Acta*, **961**: 299-308.

Bläckberg, L., and Hernell, O. (1981). *Eur. J. Biochem.*, **116**: 221-225.

Bollag, D. M., and Edelstein, S. J. (1991). *Protein methods*. New York, Wiley-Liss.

Calame, K. B., Gallo, L., Cheriathundam, E., Vahouny, G. V., and Treadwell, C. R. (1975). *Arch Biochem. Biophys.*, **168:** 57-65.

Cox, D. G., Leung, C. K. T., Kyger, E. M., Silburg, C. A., Lange, L. G. (1990). *Biochemistry*, **29:** 3842-3848.

Cuatrecasas, P., Wilchek, M., and Anfinsen, C. B. (1968). *Proc. Nat. Acad. Sci. USA*, **61:** 636-643.

Erlanson, C. (1975). *Scand. J. Gastroenterol.*, **10:** 401-408.

Hames, B. D. (1990). In *Gel electrophoresis of proteins: a practical approach*. pp. 1-139. Oxford, IRL Press.

Hernell, O., and Olivecrona, T., (1974). *Biochim. Biophys. Acta*, **369:** 234-244.

Hyun, J., Kothari, H., Herm, E., Mortenson, J., Treadwell, C. R., and Vahouny, G. V. (1969). *J. Biol. Chem.*, **214:** 1937-1945.

Hyun, J., Steinberg, M., Treadwell, C. R., and Vahouny, G. V. (1971). *Biochem. Biophys. Res. Comm.*, **44:** 819-825.

Hyun, J., Treadwell, C. R., and Vahouny, G. V. (1972). *Arch. Biochem. Biophys.*, **152:** 233-242.

Labow, R., Adams, K., and Lynn, K. R. (1983). *Biochim. Biophys. Acta*, **749:** 32-41.

Lombardo, D., Guy, O., and Figarella, C. (1978). *Biochim. Biophys. Acta*, **527:** 142-149.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951). *J. Biol. Chem.*, **193:** 265-275.

Momsen, W. E., and Brockman, H. L. (1977). *Biochim. Biophys. Acta*, **486:** 103-113.

Moore, A., Burton, G. W., and Ingold, K. U. (1994). *J. Am. Chem. Soc.*, **116:** 6945-6946.

Riley, D. J. S., Kyger, E. M., Spilburg, C. A., and Lange, L. G. (1990). *Biochemistry*, **29:** 848-3852.

Roberts, I. M., Wiesenfeld, P., and Jacobson, P. (1988). *LC-GC*, **6:** 144-146.

Swan, J. S., Hoffman, M. M., Lord, M. K., and Poechmann, J. L. (1992). *Biochem. J.*, **283:** 119-122.

Vahouny, G. V., Weersing, S., and Treadwell, C. R. (1965). *Biochem. Biophys. Acta*, **98:** 607-616.

Wang, C.-S. (1980). *Anal. Biochem.*, **105:** 398-402.

5B. References

PART B. Taurocholate-Induced Dimerization of Bovine Cholesterol Esterase in Sodium Dodecylsulfate

Bayreuther, K., Beiseler, B., Ehring, R., Griesser, H.W., Meischendahl, M., Muller-Hill, B. and Triesch, I. (1980) *Biochem. Soc. Trans.*, **8:** 675-676.

Bensadoun, A., Ehnholm, C., Steinberg, D. and Brown, W.V. (1974) *J. Biol. Chem.*, **249:** 2220-2227.

Bollag, D.M. and Edelstein, S.J. (1991) *Protein Methods*, p. 100, Wiley-Liss, New York.

Calame, K.B., Gallo, L., Cheriathundam, E., Vahouny, G.V. and Treadwell, C.R. (1975) *Arch. Biochem. Biophys.*, **168:** 57-65.

Clarke, D.M. and Bragg, P.D. (1985) *J. Bacteriol.*, **162:** 367-373.

Cox, D.G., Leung, C.K.T., Kyger, E.M., Spilburg, C.A. and Lange, L.G. (1990) *Biochemistry*, **29:** 3842-3848.

Deutsch, D.G. (1976) *Anal. Biochem.*, **71**: 300-303.

Erlanson, C. (1975) *Scand. J. Gastroenterol.*, **10**: 401-408.

Gallagher, S.R. and Leonard, R.T. (1987) *Anal. Biochem.*, **162**: 350-357.

Hames, B.D. (1990) in *Gel Electrophoresis of Proteins: A Practical Approach*, (Hames, B.D. and Rickwood, D., Eds.), pp. 1-139, IRL Press, Oxford.

Hyun, J., Kothari, H., Herm, E., Mortenson, J., Treadwell, C.R. and Vahouny, G.V. (1969) *J. Biol. Chem.*, **244**: 1937-1945.

Hyun, J., Steinberg, M., Treadwell, C.R. and Vahouny, G.V. (1971) *Biochem. Biophys. Res. Commun.*, **44**: 819-825.

Hyun, J., Treadwell, C.R. and Vahouny, G.V. (1972) *Arch. Biochem. Biophys.*, **152**: 233-242.

Jacobson, P.W., Wiesenfeld, P.W., Gallo, L.L., Tate, R.L. and Osborne, J.C., Jr. (1989) *J. Biol. Chem.*, **265**: 515-521.

Larson, T.J., Schumacker, G. and Boos, W. (1983) *J. Bacteriol.*, **152**: 1008-1021.

Leach, B.S., Collawn, J.F. and Fish, W.W. (1980) *Biochemistry*, **19**: 5734-5741.

Lin, J.-L., Wang, W.-Y., Liao, T.-H.(1994) *Biochim. Biophys. Acta*, **1209**: 209-214.

Lombardo, D., Guy, O. and Figarella, C. (1978) *Biochim. Biophys. Acta*, **527**: 142-149.

Meyer, J. G. (1989) *Biochim. Biophys. Acta*, **1002**: 89-92.

Moore, A.N.J., Vakos, H.T, Neville, T.A.M., Tonge, P., Arya, P., and Burton, G.W. (1996) *Biotechnology Techniques* **10**: 523-528.

Rudd, E. A., Mizuno, N. K., and Brockman, H. L. (1987) *Biochim. Biophys. Acta*, **918**: 106-114.

See, Y.P. and Jackowski, G. (1989) in *Protein Structure: A Practical Approach* (Creighton, T.E., Ed.), pp. 1-22, IRL Press, Oxford.

Stocker, N.G., Pratt, J.M. and Spratt, B.G. (1983) *J. Bacteriol.*, **155**: 854-859.

Talkington, D.F., Voellinger, D.C., McDaniel, L.S. and Briles, D.E. (1992) *Microb. Pathogen.*, **13**: 343-355.

Teather, R.M., Muller-Hill, B., Abrutsch, U., Aichele, G. and Overath, P. (1978) *Mol. Gen. Genet.*, **159**: 239-248.

Vahouny, G.V., Weersing, S. and Treadwell, C.R. (1965) *Biochim. Biophys. Acta*, **98**: 607-616.

Vahouny, G.V. and Brecher, A.S. (1968) *Arch. Biochem. Biophys.*, **123**: 247-254.

Vakos, H. T., Burton, G. W., and Kaplan, H. (1997) *Biochim. Biophys. Acta*, **1342**: 103-108.

5B.1 References

PART B. Prospective Studies for Cholesterol Esterase: Use of Low Water Reverse Micellar Systems to Probe Structure-Function Relationships

Barbaric, X. and Luisi, P. L. (1981) *J. Am. Chem. Soc.*, **103**: 4239-4244.

Barrabin, H., Scofano, H. M., Tuena de Gómez-Puyou and Gómez-Puyou, A. (1993) *Eur. J. Biochem.*, **213**: 757-763.

Bru, R. and Garcia-Carmona, F. (1991) *FEBS Lett.*, **282**: 170-174.

Bru, R., Sánchez-Ferrer, A., and García-Carmona, F. (1995) *Biochem J.*, **310:** 721-739.

Bru, R. and Walde, P. (1991) *Eur. J. Biochem.*, **199:** 95-103.

Bru, R. and Walde, P. (1993) *Biochem. Mol. Biol. Int.*, **31:** 685-692.

Chang, G.-G. and Shiao, S.-L. (1994) *Eur. J. Biochem.*, **220:** 861-870.

Darszon, A. and Shoshani, L. (1992) *Enzymes in Reverse Micelles Containing Phospholipids in Biomolecules in Organic Solvents*, Gómez-Puyou, A., Darszon, A., and Tuena de Gómez-Puyou, M., eds., CRC Press, Boca Raton, FL. p. 33-65.

Davis, D. M., Mcloskey, D., Birch, D. J. S., Gellert, P. R., Kittlety, R. S., and Swart, R. M. (1996) *Biophys. Chem.*, **60:** 63-77.

De Meis, L. Martins, O. B., and Alves, E. W. (1980) *Biochemistry*, **19:** 4252-4261.

Desfosses, B. Cittanova, N., Urbach, W., and Waks, M. (1991) *Eur. J. Biochem.*, **199:** 79-87.

Desfosses, B., Nicot, C., and Waks, M. (1992) *Biochem. Int.*, **26:** 257-264.

Dorovska-Taran, V. N., Veeger, C., and Visser, A. J. W. G. (1993) *Eur. J. Biochem.*, **211:** 47-55.

Escamilla, E., Ayala, G., Tuena de Gómez-Puyou, M., Gómez-Puyou, A., Millán, L., and Darszon, A. (1989) *Arch. Biochem. Biophys.*, **272:** 332-343.

Fadnavis, N. W., Chandraprakash, Y., and Deshpande, A. (1993) *Biochimie*, **75:** 995-999.

Fernández-Velasco, D. A., Garza-Ramos, G., Ramirez, L., Shoshani, L., Darszon, A., Tuena de Gómez-Puyou, M., and Gómez-Puyou, A. (1992) *Eur. J. Biochem.*, **205:** 501-508.

Fernández-Velasco, D. A., Sepulveda-Becerra, M., Galina, A., Darszon, A., Tuena de Gómez-Puyou, M., and Gómez-Puyou, A. (1995) *Biochemistry*, **34:** 361-369.

Fletcher, P. D., Howe, A. M., and Robinson, B. H. (1987) *J. Chem. Soc. Faraday Trans. I.*, **87:** 985-1006.

Gallay, J., Vincent, M., Nicot, C., and Waks, M. (1987) *Biochemistry*, **26:** 5738-5747.

Gardell, S. J. and Tate, S. S. (1981) *J. Biol. Chem.*, **256:** 4799-4804.

Garza-Ramos, G., Darszon, A., Tuena de Gómez-Puyou and Gómez-Puyou, A. (1992) *Biochemistry*, **29:** 751-757.

Garza-Ramos, G., Tuena de Gomez-Puyou, Gomez-Puyou, A., and Gracy, R. W. (1992) *Eur. J. Biochem.*, **208:** 389-395.

Gonnelli, M. and Strambani, G. B. (1988) *J. Phys. Chem.*, **92:** 2854-2857.

Green, S. M., Gittis, A. G., Meeker, A. K., and Lattman, E. E. (1995) *Nature New Biol.*, **2:** 746-751.

Hanley, A. B., Furniss, C. S. M., Kwiatkowska, C. A., and Mackie, A. R. (1991) *Biochim. Biophys. Acta*, **1074:** 40-44.

Hauser, H., Haering, G., Pande, A., and Luisi, P. L. (1989) *J. Phys. Chem.*, **93:** 7869-7876.

Jaenicke, R., Vogel, W., and Rudolph, R. (1981) *Eur. J. Biochem.*, **114:** 525-531.

Kabanov, A. V., Nametkin, S. N., Evtushenko, G. N., Chernov, N. N., Klyachko, N. L., Levashov, A. V., and Marinek, K. (1989) *Biochim. Biophys. Acta*, **996:** 147-152.

Kabanov, A. V., Nametkin, S. N., Klyachko, N. L., and Levashov, A. V. (1990) *FEBS Lett.*, **267:** 236-238.

Kabanov, A. V., Nametkin, S. N., Klyachko, N. L., and Levashov, A. V. (1991) *FEBS Lett.*, **278:** 143-146.

Katiyar, S. S., Kumar, A., and Kumar, A. (1989) *Biochem. Int.*, **19:** 547-552.

Kernen, P., Agositi, R. D., Strasser, R. J., and Darszon, A. (1997) *Biochim. Biophys. Acta*

Khmelnitsky, Y. L., Kabanov, A. V. Klyachko, N. L., Levashov, A. V., and Martinek, K. (1989) *Enzyme Catalysis in Reverse Micelles*, Pileni, M. P., ed., Elsevier, Amsterdam.

Khmelnitsky, Y. L., Hilhorst, R., Visser, A. J. W. G. and Veeger, C. (1993a) *Eur. J. Biochem.*, **211:** 73-77.

Khmeltnitsky, Y. L., Hoek, A., Veeger, C. and Visser, A. J. W. G. (1993b) *Eur. J. Biochem.*, **212:** 63-67.

Klyachko, N. L., Levashov, A. V., Pshezhetsky, A. V., Bogdanova, N. G., Berezin, I. V., and Martinek, K. (1986) *Eur. J. Biochem.*, **161:** 149-154.

Lalitha, J. and Mulimani, V. H. (1996) *Biochem. Mol. Biol. Int.*, **40:** 571-578.

Lenz, V. J., Federwisch, M., Gattner, Hans-Gregor, Brandenburg, D., Höcker, H., Hassiepen, U., and Wollmer, A. (1995) *Biochemistry*, **34:** 6130-6141.

Luisi, P. L., Giomini, M., Pileni, M. P., and Robinson, B. H. (1988) *Biochim. Biophys. Acta*, **947:** 209-246.

Luisi, P. L., Häring, G., Maestro, M., and Rialdi, G. (1990) *Thermochim. Acta*, **162:** 1-16.

Mao, Q. and Walde, P. (1991) *Biochem. Biophys. Res. Commun.*, **178:** 1105-1112.

Martinek, K., Levashov, A. V., Khmelnitsky, Y. L., Klyachko, N. L., and Berezin, I. V. (1982) *Science*, **218:** 889-891.

Martinek, K. (1989) *Biochem. Int.*, **18:** 871-893.

Marzola, P. and Gratton, E. (1991) *J. Phys. Chem.*, **95:** 9488-9495.

McCracken, S. and Meighen, E. (1980) *J. Biol. Chem.*, **255:** 2396-2404.

Menger, F. M. and Yamada, K. (1979) *J. Am. Chem. Soc.*, **101:** 6731-6734.

Nicot, C. and Waks, M. (1996) *Biotechnol. Gen. Engin. Rev.*, **13:** 267-314.

Qinglong, C., Huizhou, L., and Jiayong, C. (1994) *Biochim. Biophys. Acta*, **1206:** 247-252.

Sepulveda-Becerra, M., Ferreira, S. T., Strasser, R. J., Garzón-Rodriguez, W., Beltrán, C., Gómez-Puyou, A. and Darszon, A. (1996) *Biochemistry*, **35:** 15915-15922.

Singh, A. K. and Aruna, R. V. (1995) *Biochim. Biophys. Acta*, **1245:** 167-172.

Shastry, M. C. R. and Eftnik, M. R. (1996) *Biochemistry*, **35:** 4094-4101.

Shoshani, L., Darszon, A., Tuena de Gómez-Puyou, M., and Gómez-Puyou, A. (1994) *Eur. J. Biochem.*, **221:** 1027-1032.

Strambini, G. B. and Gonnelli, M. (1988) *J. Phys. Chem.*, **92:** 2850-2853.

Tang, S.-S. and Chang, G.-G. (1996) *Biochem. J.*, **315:** 599-606.

Thanki, N., Zeelen, J. P., Mathieu, M., Jaenicke, R., Abagyan, R. A., Wierenga, R. K., and Schlieb, W. (1997) *Prot. Eng.*, **10:** 159-167.

Verhaert, R. M. D., Hilhorst, R., Visser, A. J. W. G., and Veeger, C. (1992) *The Optimization of Enzyme Catalysis in Organic Media of Biomolecules in Organic Solvents*, Gómez-Puyou, A., Darszon, A., and Tuena de Gómez-Puyou, M., eds., CRC Press, Boca Raton, FL, pp.133-162.

Visser, A. J. W. G., van Engelen, J., Visser, N. V., van Hoek, A., Hilhorst, R., and Freedman, R. B. (1994) *Biochim. Biophys. Acta*, **1204:** 225-234.

Voss, K., Laane, C., and Visser, A. J. W. G. (1987) *Photochem. Photobiol.*, **45:** 863-878.

Walde, P., Mao, Q., Bru, R., Luisi, P. L., and Kuboi, R. (1992) *Pure Appl. Chem.*, **64:** 1771-1775.

Waley, S. G. (1973) *Biochem. J.*, **135:** 165-172.

Wang, X., Wang, C. S., Tang, J., Dyda, F., and Zhang, X. C. (1997) *Structure*, **5:** 1209-1218.

Zabori, S., Rudolph, R., and Jaenicke, R. (1980) *Z. Naturforsch.*, **35c:** 999-1004.

Zulauf, M. and Eicke, H. F. (1979) *J. Phys. Chem.*, **83:** 480-486.

5C. References

PART C. Bile Salt-Modulated Stereoselection in the Cholesterol Esterase-Catalyzed Hydrolysis of α-Tocopheryl Acetates

Atkins, W. M., Sligar, S. G. (1989) *J. Am. Chem. Soc.*, **111:** 2715-2717.

Bläckberg, L., Hernell, O. (1993) *FEBS Lett.*, **323:** 207-210.

Bläckberg, L.; Lombardo, D.; Hernell, O.; Guy, O.; Olivecrona, T. (1981) *FEBS Lett.*, **136:** 284-288.

Blow, D. M. (1976) *Acc. Chem. Res.*, **9:** 145-152.

Borgström, B. J. (1967) *J. Lipid Res.*, **8:** 598-608.

Calame, K. B.; Gallo, L.; Cheriathundam, E.; Vahouny, G. V.; Treadwell, C. R. (1975) *Arch. Biochem. Biophys.*, **168:** 57-65.

Carey, M. C.; Small, D. M. (1972) *Arch. Intern. Med.*, **130:** 506-527.

Carey, M. C.; Small, D. M. (1970) *Am. J. of Med.*, pp. 500-608.

DiPersio, L. P., Fontaine, R. N., Hui, D. Y. (1991) *J. Biol. Chem.*, **266:** 4033-4036.

Erlanson, C. (1975) *Scand. J. Gastroenterology*, **10:** 401-408.

Fersht, A. (1985) Enzyme Structure and Mechanism, 2nd. Ed., Freeman, New York, Ch. 8.

Fersht, A. R., Shi, J.-P., Wilkinson, A. J., Blow, D. M., Carter, P., Waye, M. M. Y., Winter, G. P. (1984) *Angew Chem., Int. Ed. Engl.*, **23:** 467-473.

Gallo-Torres, H. E. (1970) *Lipids*, **5:** 379-384.

Gallo-Torres, H. E. (1970) *Int. J. Vit. Res.*, **40:** 505-514.

Helenius, A.; McCaslin, D. R.; Fries, E.; Tanford, C. (1979) *Methods Enzymol.*, **56:** 734-749.

Hughes, L.; Slaby, M.; Burton, G. W.; Ingold, K. U. (1990) *J. Labeled Comp. Radiopharmacol.*, **28:** 1049-1057.

Hyun, J.; Kothari, H.; Herm, E; Mortenson, J.; Treadwell, C. R.; Vahouny, G. V. (1969) *J. Biol. Chem.*, **244:** 1937-1945.

Hyun, J.; Treadwell, C. R.; Vahouny, G. V. (1972) *Arch. Biochem. Biophys.*, **152:** 233-242.

Ingold, K. U., Burton, G. W., Foster, D. O., Hughes, L., Lindsay, D. A., Webb, A. (1987a) *Lipids*, **22:** 163-172.

Ingold, K. U.; Hughes, L.; Slaby, M.; Burton, G. W. (1987b) *J. Labeled Comp. Radiopharmacol.*, **24:** 817-831.

Knowles, J. R. (1987) *Science*, **236**: 1252-1258.

Korzenovsky, M., Diller, E. R., Marshall, A. C., Auda, B. M. (1960) *Biochem. J.*, **76**: 238-245.

Lombardo, D. (1982) *Biochim. Biophys. Acta*, **700**: 67-74.

Lombardo, D., Campese, D., Multigner, L., Lafont, H., De Caro, A. (1983) *Eur. J. Biochem.*, **133**: 327-333.

Lombardo, D.; Deprez, P.; Guy, O. (1980a) *Biochimie*, **62**: 427-432.

Lombardo, D.; Fauvel, J.; Guy, O. (1980b) *Biochim. Biophys. Acta*, **611**: 136-146.

Lombardo, D.; Fauvel, J.; Guy, O. (1980) *Biochim. Biophys. Acta*, **611**: 147-155.

Lombardo, D. and Guy, O. (1979) *Biochimie*, **61**: 415-417.

Lombardo, D. and Guy, O. (1980) *Biochim. Biophys. Acta*, **611**: 415-417.

Lombardo, D. and Guy, O. (1981) *Biochim. Biophys. Acta*, **657**: 425-437.

Lynn, K. R. Chuaqui, C. A., Clevette-Radford, N. A. (1982) *Biorg. Chem.*, **11**: 19-23.

March, J. (1985) In *Advanced Organic Chemistry*, Third Edition, Wiley-Interscience, New York. (a) 334-338; (b) 349-350.

Mathias, P. M.; Harries, J. T.; Muller, D. P. R. (1981a) *J. Lipid Res.*, **22**: 177-184.

Mathias, P. M.; Harries, J. T.; Peters, T. J.; Muller, D. P. R. (1981b) *J. Lipid Res.*, **22**: 829-837.

Moore, A., Burton, G. W., Ingold, K. U. (1994) *J. Am. Chem. Soc.*, **116**: 6945-6946.

Moore, A. N. J., Dutton, P. J., Zahalka, H. A., Burton, G. W., and Ingold, K. U. (1995) *J. Am. Chem. Soc.*, **117**: 5677-5686.

Muller, D. P. R.; Manning, J. A.; Mathias, P. M.; Harries, J. T. (1976) *Int. J. Vit. Nutr. Res.*, **46**: 207-211.

Norman, A. (1954) *Acta Physiol. Scand.*, **32**: 1-10.

Ransac, S.; Riviere, C.; Soulie, J. M.; Gancet, C.; Verger, R.; de Haas, G. H. (1990) *Biochim. Biophys. Acta*, **1043**: 57-66.

Rudd, E. A. and Brockman, H. L. (1984) In: *Lipases*, Borgstrom, B. and Brockman, H. L. (Editors), Part II, Ch. D, Elsevier Press, New York, Oxford, pp. 185-204.

Still, W. C.; Kahn, M.; Mitra, A. (1978) *J. Org. Chem.*, **43**: 2923-2925.

Stout, J. S., Sutton, L. D., Quinn, D. M. (1985) *Biochim. Biophys. Acta*, **837**: 6-12.

Stroud, R. M. (1974) *Sci. Am.*, **231**: 74-88.

Subbiah, M. T., Kuksis, A., Mookerjea, S. (1969) *Can. J. Biochem.*, **47**: 847-854.

Sutton, L. D.; Stout, J. S.; Quinn, D. M. (1990) *J. Am. Chem. Soc.*, **112**: 8398-8403.

Tsujita, T., Okuda, H. (1990) *J. Lipid Res.*, **31**: 831-838.

Vahouny, G. V.; Weersing, S.; Treadwell, C. R. (1964) *Arch. Biochem. Biophys.*, **107**: 7-15.

Vahouny, G. V.; Weersing, S.; Treadwell, C. R. (1965) *Biochim. Biophys. Acta*, **98**: 607-616.

Weiser, H. and Vecchi, M. (1981) *Internat. J. Vit. Nutr. Res.*, **51**: 100-113.

Wescott, C.R. and Klibanov, A.M. (1994) *Biochim. Biophys. Acta*, **1206**: 1-9.

Wang, C.-S., Hartsuck, J. A. (1993) *Biochim. Biophys. Acta*, **1166**: 1-19.

Wu, S.-H., Guo, Z.-W. and Sih, C. J. (1990) *J. Am. Chem. Soc.*, **112**: 1990-1995.

Yousef, I. M., Kakis, G., Fisher, M. M. (1972) *Can. J. Biochem.*, **50**: 402-408.

Zahalka, H. A.; Cheng, S. C.; Burton, G. W.; Ingold, K. U. (1987) *Biochim. Biophys. Acta*, **921**: 481-485.

Zahalka, H. A., Dutton, P. J., O'Doherty, B., Smart, T. A.-M., Phipps, J., Foster, D. O., Burton, G. W. and Ingold, K. U. (1991) *J. Am. Chem. Soc.*, **113**: 2797-2799.

Concluding Remarks

The field of protein chemistry utilizes both chemical and physical approaches in an overall strategy to solve a research problem (Chapter 1). The theme of the present work was to use the *in vacuo* chemical modification approach of Taralp and Kaplan (1997), coupled with physical techniques (i.e., NMR spectroscopy, titration, SDS-PAGE, autoradiography) and activity assays, to elucidate protein structure and function. Taralp (1997, Doctoral Thesis) has shown that a nonaqueous environment provides the necessary framework to carry out useful chemical modifications. Practical applications have been developed and are demonstrated in the present research.

NMR spectroscopy was used as an analytical tool for the identification and characterization of N- and C-termini of protein primary structure (Chapters 2-4). Apart from the known chemical methods, there is no facile *physical* technique at present that can be used to identify the amino- or carboxyl-terminus of a protein or to determine the presence of multiple amino- or carboxyl-termini in a protein preparation. Current NMR technology, coupled with the nonaqueous *in vacuo* chemical modification technique (Taralp and Kaplan, 1997), have allowed for the development of sensitive methods for investigating protein purity, identity, and homogeneity. The *in vacuo* chemical modification procedure makes it possible to incorporate labeled reagents (^{13}C, ^{14}C, ^{2}H, ^{3}H) into protein with a high degree of reaction specificity in a cost-effective manner, since only microlitre quantities are required to modify milligram quantities of lyophilized protein.

In Chapters 2 and 4, the *in vacuo* chemical modification technique was applied to N-terminal determination. Tests for characterizing protein integrity are especially important in the biotechnology industry, where new biotechnological products such as recombinant

proteins intended for therapeutic purposes must be of the utmost purity. The pH memory effect (Zaks and Klibanov, 1988) was used to control the ionization state of amino groups in lyophilized proteins, and hence, their chemical reactivity towards modifying reagents. Lyophilized proteins were reacted *in vacuo* (at 75°C) with ^{13}C-iodomethane at LpH 6-7 to yield preferentially or selectively ^{13}C-trimethylated α-amino groups, which were observed by ^{13}C-NMR spectroscopy. In theory, the chemical shift of the ^{13}C-trimethylamino derivative of the N-terminal amino acid derived from ^{13}C-methylated protein via acid hydrolysis is more distinct than the same derivative in the intact protein and this was the original route of investigation. However, the problem of incomplete hydrolysis arose and the investigation reverted to the utilization of the intact ^{13}C-methylated protein derivatives (Chapter 2). Adequate sensitivity allowed for the detection of minor resonances of free α-amino groups arising from proteolysis of the major protein or from protein impurities. It is particularly noteworthy that SDS-PAGE cannot distinguish subtle differences of contaminating proteins of similar size or heterogeneity due to proteolysis during the purification or isolation process. The resonances of the ^{13}C-trimethylated α-amino groups in standard amino acids and peptides are sufficiently close to those in the ^{13}C-methylated protein to make a tentative identification of the *number and types* of N-terminal amino acid residues. It was also possible to verify whether a protein has a *free or blocked* amino-terminus. *In vacuo* reaction with ^{13}C-iodomethane also resulted in the formation of ^{13}C-methyl esters of the carboxyl groups in the lyophilized protein. The success of the N-terminal identification methodology was dependent on the identification and removal (with hydroxylamine or base and heat treatments) of these resonances from the ^{13}C-NMR spectral region between 52 and 57 ppm, as they overlapped some of the trimethylamino resonances. Simplification of the NMR

spectra made it possible to unequivocally identify resonances arising from trimethylamino groups.

Both ^{13}C and ^{14}N NMR spectroscopy (Chapters 2 and 4) were utilized to characterize the trimethylammonium derivatives of α- and ε-amino groups formed by reaction with iodomethane (^{13}C, ^{12}C). The use of ^{14}N NMR spectroscopy has been limited with regard to the study of proteins. Iodomethane affords a unique opportunity to directly observe and investigate the reaction, the local environment, and motion of amino groups in any protein. The ^{14}N NMR spectroscopic method for identifying trimethylated amino groups required that the protein be hydrolyzed (Chapter 4). Standard compounds (i.e., methylated amino acids) were prepared in two ways, by *in vacuo* methylation and by exhaustive methylation; the latter known protocol corroborated the former new approach. The *in vacuo* approach utilized very small quantities of amino acid starting material (1-2 mg) compared to the several milligrams required in the protocol for exhaustive methylation. The ^{14}N NMR spectroscopic approach to N-terminal identification is good from a theoretical point of view, but clearly is inferior to the ^{13}C-NMR spectroscopic approach due to the sensitivity problem and the requirement of long acquisition times and large amounts of protein. If the latter is not of concern, however, the method has the advantage of being cost- effective because ^{14}N is 99.6% naturally abundant. Moreover, it can be utilized to trimethylate any amino group in the protein molecule (i.e., via trace labeling) and dynamic studies can yield information on the motion of the amino group in its local environment in the protein structure, with the possibility of offering functional information. It is noteworthy that the ^{13}C NMR spectroscopic approach does not necessitate that the protein be hydrolyzed, as does the ^{14}N NMR method, which is a direct consequence of the nuclear properties of carbon-13 and

nitrogen-14—the former has a spin of ½ and has a spherical nuclear charge distribution, and the latter has a spin of 1 and is quadrupolar (nuclear charge distribution deviates from spherical symmetry) with extremely efficient relaxation properties in the presence of a magnetic field, and hence, possesses different NMR properties than carbon-13, namely broader lines unless the electronic distribution around the ^{14}N nucleus is made more symmetrical, as in a trimethylammonium derivative of an amino acid (Chapter 4).

The utility of iodomethane as a protein-modifying reagent (Taralp and Kaplan, 1997) was further exemplified in the present research (Chapter 3). The pH memory effect (Zaks and Klibanov, 1988) was confirmed by the *in vacuo* reaction of lyophilized proteins with ^{13}C-iodomethane at low LpH values (between 3-4), which effected preferential esterification of carboxyl groups. An increase in the extent of methylation was observed with an increase of deprotonated carboxyl groups (i.e., with increasing LpH values). Trimethylation of amino groups was the predominant reaction at LpH values greater than 4 or 5. In contrast, the novel *in vacuo* esterification reaction on lyophilized protein (LpH 3) with gaseous ^{13}C-methanol/HCl or ^{13}C-ethanol/HCl at 75°C resulted in the exclusive formation of α-, γ-, and δ-carboxyl ^{13}C-methyl or ^{13}C-ethyl ester derivatives, which were distinguishable by their distinct chemical shifts. Hence, *in vacuo* esterification with alcohol/HCl is the obvious procedure to use for specific methylation of lyophilized protein carboxyl groups. Although 100% methyl-esterification of carboxyl groups is difficult to achieve, the extent attainable (60-70%) is as good as the conventional method for acid-catalyzed esterification by suspension of the protein in acidic alcohol (Ram and Maurer, 1959; Broomfield *et al.*, 1965; Means and Feeney, 1971). Hence if extensive esterification of carboxyl groups is desired, then *in vacuo* reaction with methanol/HCl is the method of choice. Ethanol/HCl esterified

carboxyl groups in the lyophilized protein to a much lesser extent (25%) than did methanol/HCl. It appears that not all carboxyl groups in the lyophilized protein are freely accessible and the bulkiness of the alcohol is an important factor in determining the extent of esterification. *In vacuo* esterification with ethanol/HCl could be very useful in cases where it is desirable to form a more stable ester derivative, or to introduce a more stable ester derivative containing a ^{14}C, ^{13}C, CD_2, CD_3, CT_2 or CT_3 label into the protein. Reaction with iodomethane was useful in that the pH memory effect of lyophilized proteins (Zaks and Klibanov, 1988; Taralp, Doctoral Thesis, 1997) was confirmed; but the results show that when alkylating amino groups (Chapter 2; Vakos *et al.*, 2000), *in vacuo*-synthesized esters of carboxyl groups will be formed and should be removed. Moreover, it is much easier to control the extent of reaction with iodomethane, so in cases where a low extent of esterification of carboxyl groups is desired, reaction with iodomethane would be preferable. The current NMR technology provided sufficient sensitivity to identify the ^{13}C-chemical shifts of the different types of methyl esters that can be formed by reacting protein with methylating reagents. Resonances from individual esterified side-chains of aspartic (γ) or glutamic (δ) acid residues could not be greatly resolved. In contrast, the chemical shift of the C-terminal α-carboxyl group showed a strong dependence on the nature of the C-terminal amino acid and C-terminal sequence. It should be possible to make a tentative identification of the C-terminal amino acid(s) of a protein by comparing the chemical shift(s) of the α-carboxyl ^{13}C-methyl ester observed in the protein with that of an acetylated dipeptide having the appropriate C-terminus.

The methodologies developed in Chapters 2 and 3 make it possible to detect contaminating proteins of similar size (which are difficult to detect by SDS-PAGE),

proteolytically processed proteins with more than two polypeptide chains, post-translationally chemically-modified proteins with a *blocked* N-terminus, and preproteins that are nonhomogeneously proteolyzed to the mature protein, which may have N- and C-terminal heterogeneities of primary structure.

The extent of derivatization of protein carboxyl groups can be controlled (extent depends on reaction time, LpH, pK_a and extent of surface exposure of functional groups) and exploited in structure-function or protein characterization studies, as the protein can retain its structural integrity under the conditions used. Control of the extent of derivatization is of the utmost importance because it allows for chemical modification without substantial perturbation of protein structure, and likely, affect on activity or function. The nonaqueous methodology can be utilized to effect site-specific or differential modification of carboxyl groups by fine-tuning the reaction conditions and exploiting the pH memory effect. Moreover, it affords the opportunity to use small amounts of reagents, which allows for the introduction of labels (^{13}C, ^{14}C, etc.) without disruption of protein structure (Taralp, Doctoral Thesis, 1997; Stewart *et al.*, 1997).

It is possible to form stable activated derivatives of protein carboxyl groups by the *in vacuo* chemical modification technique, which can subsequently be reacted with nucleophiles in coupling reactions, with relatively small amounts of reagents cf. the high concentrations of activating reagents and nucleophiles that are required for aqueous modifications. In this manner the physicochemical properties of a protein can be altered [e.g., a negatively charged carboxyl group can be converted to a positively charged functionality by reacting its activated form with a diamine, followed by reaction with iodomethane to form the trimethylammonium salt of the group (Taralp, 1997, Doctoral

Thesis)]. Since the *in vacuo* approach to chemical modification of protein allows for the formation of derivatives with different degrees of modification, sequential modifications can afford new insights into protein structure and function. However, a high extent of modification may be desired in certain applications if maintenance of activity of the protein is not a concern. Significantly, the development of this methodology (i.e., carboxyl group methylation or esterification) has led to a working strategy for a C-terminal identification protocol [refinements made by Mr. Nicolas Stewart (Ph.D. thesis to be submitted)].

Another important finding is the results observed with methylated human serum albumin at low LpH values; there was significant reaction with Lys ε-amino groups with iodomethane. This result is significant and indicated that proton transfer may be occurring between a protonated amino group and an anionic counterion, and suggests that a dynamic equilibrium between ionization states in the lyophilized protein exists, but is very slow, and that under the *in vacuo* conditions employed, small amounts of unprotonated nucleophilic amino groups are free to react with electrophilic iodomethane. Evidence for protein crosslinking was also observed via SDS-PAGE with esterified ribonuclease A (at low LpH). A suggested mechanism is the displacement of methanol by the reaction of an amino group with a carboxyl methyl ester; proton transfer may be involved in this mechanism since the dimerization occurred in protein derivatized at low LpH values, under which conditions amino groups are protonated. If proton transfer occurs under the *in vacuo* conditions, it is not unreasonable to suggest that a crosslink may form at an intermolecular salt bridge with the concomitant loss of water. This avenue of research is currently being pursued by Ms. Brigitte Simons.

Implementation of the *in vacuo* chemical modification technique of Taralp and Kaplan (1997) has allowed for the preparation of radiolabeled protein molecules with high specific activity using ^{14}C-iodomethane for the purpose of elucidating their mechanisms of action in tracer studies. Such high levels of specific activity cannot be achieved in a cost efficient manner if the radiolabeling is carried out in aqueous environments due to the competition of water with the protein for the reagent, which can be very expensive. Derivatization of lyophilized proteins under nonaqueous conditions can maintain the essential elements of the native protein structure and can also avoid some of the difficulties encountered with aqueous modification (i.e., proteolysis, denaturation). Proteins that are not enzymes can be chemically tagged in a manner that can serve as a probe of structure and function. In the present study, insulin and *Bacillus thuringiensis* (*Bt*) insecticidal toxin were trace-radiolabeled. ^{14}C-Methylated *Bt* toxin was shown to retain partial activity. Further studies are required, but the preliminary results were promising and foreshadow the widespread use of *in vacuo* derivatizations for trace-labeling protein for use in tracer studies.

The investigations of this study have provided further supporting evidence for the general applicability of the *in vacuo* nonaqueous methodology to effect the chemical modification of proteins for the purpose of studying structure-function relationships.

Chapter 6 dealt with a separate study of the enzyme, cholesterol esterase (CE). Bovine and porcine CE (BCE, PCE) and human milk bile salt-activated lipase (BAL) were purified successfully using a cholate-derivatized affinity column and a taurocholate concentration gradient in the eluent (Moore *et al.*, 1996). The latter is an improvement on the established technology (Wang, 1980) and yields an enhanced purification compared to that attainable by the commercially available affinity resin from Sigma. SDS-PAGE was

used to study protein-protein interactions; taurocholate was shown to induce dimerization of BCE in SDS. It was shown that the conditions of sample preparation prior to performing SDS-PAGE can have unexpected effects on the apparent homogeneity of a protein. In cases where multiple bands are observed on SDS-PAGE, care should be taken to verify that the bands observed represent the individual monomeric components of the protein and not associated forms.

Protein-protein interactions can be studied in low water reverse micellar systems; and prospective studies were highlighted. Low water reverse micellar systems could provide a better understanding of the structural arrangements that monomers undergo after association to yield the active oligomer. Insight may be gained into the macromolecular interactions that occur when 3α,7α,12α-trihydroxy bile salt molecules interact with CE to induce an oligomeric active enzyme. Complex macromolecular associations have been explored, and it may be possible to stabilize different intermediate conformers of CE at different levels of association by adjustments in water concentration, and to study the acquisition of the catalytically active three-dimensional structure of CE in "slow motion" (e.g., via time-resolved fluorescence spectroscopy). It would be interesting to observe whether or not a CE monomer is active in a low water reverse micellar system, as this would be in conflict with the data in standard aqueous media.

Studies on the bile salt-modulated stereoselectivity of CE-catalyzed hydrolysis of α-tocopheryl acetates revealed that, in competitive experiments, it is possible to control the diastereoselectivity of CE by simply changing the bile salt cofactor; this is relevant if the bile salt-modulated effect on CE-catalyzed reactions is to be exploited in organic syntheses in biotechnological applications. Kinetic studies were carried out with synthesized

α-tocopheryl esters of varying chain length. The pentanoate ester showed the faster rate of hydrolysis cf. acetate and decanoate esters. The results were consistent with a rate-limiting deacylation stage of the CE mechanism for the hydrolysis reaction.

Low water systems are proving to be invaluable in advancing our knowledge and understanding of protein structure and function (Chapters 1-6). The *in vacuo* chemical modification technique clearly affords useful chemical modifications and is indispensable in an overall strategy for the development of theoretical studies and novel practical applications.

Claims to Original Research

First demonstration that the *in vacuo* reaction of lyophilized protein with ^{13}C-iodomethane at low LpH values results in the preferential esterification of carboxyl groups, as evidenced by ^{13}C-NMR spectroscopy. The observation confirms the pH memory effect of lyophilized proteins postulated by Zaks and Klibanov (1988) and verified by Taralp and Kaplan (1997). The extent of derivatization can be controlled and exploited to probe protein structure and function. The results lend further supporting evidence for the utility of ^{13}C-iodomethane as a protein-modifying reagent using the *in vacuo* methylation technique (Taralp and Kaplan, 1997).

First demonstration of the *in vacuo* reaction of protein carboxyl groups with gaseous ^{13}C-methanol or ^{13}C-ethanol using hydrogen chloride as catalyst, as evidenced by ^{13}C-NMR spectroscopy. The reaction resulted in the exclusive formation of α-, γ- and δ-carboxyl ^{13}C-methyl or ^{13}C-ethyl ester derivatives, which were distinguishable by their distinct chemical shifts. Resonances from individual esterified side-chains of aspartic (γ) or glutamic (δ) acid residues could not be greatly resolved. In contrast, the chemical shift of the C-terminal α-carboxyl group showed a strong dependence on the nature of the C-terminal amino acid and C-terminal sequence. It was possible to determine the *number* of C-termini and make a tentative identification of the C-terminal amino acid for model proteins. The extent of derivatization of protein carboxyl groups can be controlled and exploited in structure-function or protein characterization studies, as the protein can retain its structural integrity under the conditions used.

First demonstration that the current NMR technology provides sufficient sensitivity for the identification of the ^{13}C-chemical shifts of the different types of methyl esters that can be formed by reacting protein with methylating reagents.

A sensitive method was developed for analyzing the N-termini of protein primary structure for the purpose of investigating the homogeneity of a protein. Improvements were made to the ^{13}C-NMR spectroscopic analytical procedure for N-terminal identification (Taralp, Doctoral Thesis, 1997). The pH memory effect (Zaks and Klibanov, 1988) was used to control the ionization state of amino groups in lyophilized proteins, and hence, their chemical reactivity towards modifying reagents. Lyophilized proteins were reacted *in vacuo* (75°C, 24 h) with ^{13}C-iodomethane at LpH 6-7; the α-amino groups were observed by ^{13}C-NMR spectroscopy to be preferentially or selectively ^{13}C-trimethylated without substantial chemical modification of ε-amino groups. Minor resonances of free α-amino groups arising from proteolysis of the major protein or from protein impurities were detectable. SDS-PAGE cannot distinguish subtle differences in contaminating proteins of similar size or heterogeneity due to proteolysis. Tentative identification of the *number and types* of N-terminal amino acid residues (cf. standard compounds), and verification of whether a protein has a *free or blocked* amino-terminus was possible. *In vacuo* reaction with ^{13}C-iodomethane also resulted in the formation of ^{13}C-methyl esters of the protein carboxyl

groups, the resonances of which were removed by hydroxylamine or base and heat to clarify the spectra. This allowed for the direct observation of ^{13}C-resonances arising from trimethylamino derivatives of α- and ε-amino acid residues in the intact protein or of free amino acids in methylated protein hydrolysates based on the ^{13}C-chemical shifts of appropriate standard compounds. Both ^{13}C and ^{14}N NMR spectroscopy were utilized to characterize the trimethylammonium derivatives of α- and ε-amino groups formed by reaction with iodomethane (^{13}C, ^{12}C).

First demonstration of high specific activity of tracer protein molecules prepared *in vacuo* with ^{14}C-iodomethane. The high levels of specific radioactivity attainable by the *in vacuo* chemical modification technique of Taralp and Kaplan (1997) cannot be achieved if the derivatization is carried out in aqueous medium due to competition of water with protein for the reagent, which can be very expensive. Proteins that are not enzymes can be chemically tagged with retention of structural integrity to serve as a probe of structure and function. Bioactivity was retained for ^{14}C-methylated insecticidal toxin from *Bacillus thuringiensis*.

Alteration of an existing protocol for purifying bile salt-activated lipases or cholesterol esterases using an in-house prepared cholate-derivatized Sepharose 4B affinity column and a taurocholate concentration gradient in the eluent improved resolution and purification compared to commercially available gel material and the use of cholate or deoxycholate as eluent. The spacer size between the matrix and the covalently linked cholate molecule affected the purification.

First demonstration that taurocholate induces dimerization of bovine cholesterol esterase in SDS as observed on SDS-PAGE. The nature of the protein-protein interactions resulting in this dimerization was characterized.

First synthesis of α-tocopheryl esters of various chain lengths for use in a mechanistic study of cholesterol esterase.

Refereed Publications Arising from this Thesis

Moore, A.N.J., Vakos, H.T, Neville, T.A.M., Tonge, P., Arya, P., and Burton, G.W. (1996) "A Quick Method for Purifying Bile Salt-Activated Lipases," *Biotechnology Techniques* **10:** 523-528.

Vakos, H. T., Burton, G. W., and Kaplan, H. (1997) "Taurocholate-Induced Dimerization of Bovine Cholesterol Esterase in Sodium Dodecylsulfate," *Biochim. Biophys. Acta*, **1342:** 103-108.

Vakos, H. T., Kaplan, H., Black, B., Dawson, B., and Hefford, M. A. (2000) "Use of the pH Memory Effect in Lyophilized Proteins to Achieve Preferential Methylation of α-Amino Groups," *J. Protein Chem.* **19:** 231-7.

Vakos, H. T., Kaplan, H., Black, B., Dawson, B., and Hefford, M. A. (2001) "*In Vacuo* Esterification of Carboxyl Groups in Lyophilized Proteins," *J. Protein Chem.* **20:** 521-531.

Scientific Presentations

Harvey Kaplan, Nicolas Stewart, **Helen T. Vakos**, Bruce Black, Brian Dawson, and Mary-Alice Hefford "Selectivity of the Reaction of Iodomethane with Lyophilized Proteins," The Twelfth Symposium of the Protein Society, San Diego, CA, July 1998.

Helen T. Vakos, Harvey Kaplan and Graham W. Burton "Unusual Dimerization of Bovine Cholesterol Esterase in SDS," The Ninth Symposium of the Protein Society, Boston, MA, July 1995.

Helen T. Vakos, Andrew N. J. Moore, Prabhat Arya, and Graham W. Burton "An Efficient Way of Purifying Esterases," The 4th Ontario-Quebec Minisymposium in Synthetic and Bioorganic Chemistry, October 1993, University of Ottawa, Ottawa, ON, Canada

P. Arya, G. W. Burton, D. A. Lindsay, U. Wronska, T. Neville, M. Watson, K. A. Harrison, J. Klippenstein, **H. T. Vakos** "Bile Salt Modulated Chiral Recognition in Enzyme Catalyzed Reactions," Fourth International Symposium on Chiral Discrimination, Montreal, Quebec, Canada, 1993.

G. W. Burton, P. Arya, U. Wronska, K.A. Harrison, J. Klippenstein, and **H. T. Vakos** "Bile Salt-Stereomodulation of Enzymatic Reactions," Bile Acids: 1993 and the Future, Sponsored by the University of Minnesota, March 1993.

You become what you believe.