Université d'Ottawa • University of Ottawa

Université d'Ottawa · University of Ottawa

FACULTÉ DES ÉTUDES SUPÉRIEURES
ET POSTDOCTORALES

FACULTY OF GRADUATE AND
POSTDOCTORAL STUDIES

STEWART, Nicolas
AUTEUR DE LA THÈSE - AUTHOR OF THESIS

Ph.D (Chemistry)
GRADE - DEGREE

Department of Chemistry
FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

In vacuo Chemical Modifications of Proteins and Peptides

Harvey Kaplan
DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

P. Berti	K. Storey
A. Fallis	N. Tanphaichitr

J.-M. De Koninck, Ph.D.
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES — SIGNATURE — DEAN OF THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

In vacuo Chemical Modifications of Proteins and Peptides

Nicolas André Stirling Stewart

Thesis submitted to the School of Graduate Studies and Research in partial fulfillment of the degree of Doctor of Philosophy in Chemistry

Ottawa-Carleton Chemistry Institute
Department of Chemistry
University of Ottawa
Ottawa, Ontario
Canada

Candidate	Supervisor
Nicolas A.S. Stewart	Prof. Harvey Kaplan

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-494-01766-X
Our file Notre référence
ISBN: 0-494-01766-X

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

A novel approach for the chemical modification of proteins and peptides was developed, namely the chemical modification under vacuum (*in vacuo*) of lyophilized proteins and peptides. This method was used with iodomethane to produce peptides with a permanent positive charge by converting their amino groups to trimethylammonium derivatives for analysis by matrix assisted laser desorption ionization mass spectrometry (MALDI MS). Peptides with such a permanent charge have a higher ionization yield under MALDI MS resulting in a dramatic increase in detection. The signal intensity of the trimethylammonium derivative in MALDI MS is much greater compared to the unmodified peptide. The trimethylammonium derivative of α-amino groups of peptides derived from test lyophilized peptides and a lyophilized protein, and an ε-amino group of a peptide derived from the solitary lysine residue of the oxidized B-chain of insulin were shown to have a large increase in signal intensity compared to the unmodified peptides. The *in vacuo* methylation procedure is readily amenable to the use of a combination of isotopes, *e.g.* $^{12}CH_3I$ and $^{13}CH_3I$ or CH_3I and CD_3I, yielding a doublet signal to discriminate between peptide and non-peptide signals in the mass spectrum providing an even greater increase in sensitivity. This strategy was employed to isolate an N-terminal peptide derived from an enzymatic digest of human hemoglobin and to facilitate the analysis of spectra. Proteins and peptides lyophilized from slightly alkaline solutions yielded the greatest amount of derivatization. However, a peptide lyophilized under acidic conditions could also be methylated to give an increased MALDI signal. In light of these results a mechanism is proposed where the vacuum facilitates the removal of gaseous counter-ions to drive the reaction to the trimethylammonium derivative.

In order to selectively isolate the C-terminal peptides from enzymatic digests of proteins, lyophilized proteins were reacted with iodomethane under conditions where reaction occurred mainly with carboxylate groups. A proof of concept has been demonstrated for a method of general applicability for the determination of the C-terminal sequences of proteins. Separation and isolation of the C-terminal peptides generated from enzymatic digests for further sequencing either by chemical means (N-terminal Edman degradation chemistry) or by tandem mass spectrometry (MS/MS) was accomplished by the use of two dimensional (2D) high voltage paper electrophoresis (HVPE). The *in vacuo* method provided a means of incorporating a large amount of radioisotopically enriched iodomethane for the proof of concept. Digestion conditions were determined using a test protein, bovine α-chymotrypsin and two proteolytic enzymes (bovine pepsin and papain from papaya) in order to generate peptides of maximum length to provide greater sequence coverage. To make this method more applicable to small amounts of protein a small electrophoresis apparatus was designed and fluorescent markers were made to avoid the need for radioisotopes in 2D HVPE.

The structure-function relationship of a lyophilized enzyme, namely bovine α-chymotrypsin (αCT), was investigated by the *in vacuo* chemical modification method. The enzyme was lyophilized with different excipients and methylated *in vacuo* with iodomethane to probe the conformation of its active site. Iodomethane reacted *in vacuo* with lyophilized αCT to give an inactive enzyme in which the active-site imidazole was dimethylated. However, αCT co-lyophilized with the competitive inhibitors, *N*-acetyl-L-tryptophan or *N*-acetyl-D-tryptophan, was fully protected from such inactivation. In contrast, indole by itself not only did not protect the lyophilized enzyme from inactivation by iodomethane, but increased the rate of inactivation. The lyoprotectants citrate or sorbitol also showed opposite effects when co-lyophilized with αCT. Citrate protected the lyophilized enzyme from inactivation, while bound sorbitol dramatically accelerated the inactivation. Imprinting of lyophilized αCT with indole or sorbitol increased the reactivity of the active-site histidine towards iodomethane. Co-lyophilization of αCT with appropriate ligands is known to increase significantly its enzymatic activity in hydrophobic organic solvents. It is proposed that this imprinting phenomenon arises because a greater proportion of the active-sites in the lyophilized enzyme are in a catalytically favourable conformation where the imidazole of His-57 is more strongly hydrogen bonded to the carboxylate of Asp-102.

Glycation of proteins, *i.e.* the nonenzymatic reaction of protein amino groups with a reducing sugar accompanied by the loss of water to produce, *via* a Schiff base intermediate, a ketoamine (Amadori product), is a reversible reaction under aqueous conditions. Therefore, to drive this reaction an excess of reducing sugar is required for a period of days. However, it is the first step in the Maillard reaction (nonenzymatic browning) which leads to advanced glycation end-products (AGEs). *In vacuo* heating at 65°C of a co-lyophilized mixture of a protein with a small amount of reducing sugar (glucose) at a pH value between 4 and 10 produced solely the protein ketoamine derivative with the loss of water to the vacuum. This *in vacuo* glycation method allowed for the efficient incorporation of isotopically and radioisotopically enriched reducing sugars for ease of characterization. Protein glycation was complete within 16 hours and, for the most part, the glycated protein was soluble and stable at neutral pH with slow degradation under acidic (pH 2) and alkaline conditions (pH 10.5). Higher ratios of sugar

to protein resulted in the collapse of the lyophilizate during heating. ^{13}C-NMR spectra show the predominant adduct to be the ketoamine (Amadori product). Quantification indicates that even at pH 4.0, where all the amino groups are protonated, greater than 50% of these groups are glycated. Therefore it appears that glycation in the lyophilized state can proceed through the proteins' protonated amino groups. This reaction provides a potentially powerful simple synthetic avenue for the chemical modification of proteins, peptides or other amino containing molecules (*e.g.* amino surface containing supports for protein immobilization) co-lyophilized with molecules containing one or more reducing sugars.

"Dr Chumley, my mother used to say to me,
"In this world, Elwood" - she always called me Elwood - she'd say
"In this world, Elwood, you must be oh, so smart or oh, so pleasant."
For years I was smart.
I recommend pleasant.
You may quote me".

James Stewart as Elwood P. Dowd in "Harvey" (1950)

Acknowledgements

First and foremost I would like to thank my supervisor, Harvey Kaplan for his sharing of ideas and wisdom on things of science and life. It has been a great privilege.

I am very grateful for having shared time in the laboratory with the following fellow students: Alpay Taralp, PhD, Helen T. Vakos, PhD, Francois R. Clairmont, PhD, Mary C. King, PhD, Pham Van Thong, PhD graduate student and technician: Barry Lorbetski.

I would like to thank the following people for their technical support; Glenn Facey at the University of Ottawa, Faculty of Science NMR Facility for solid state NMR. Christin T. Choma at Rensselaer Polytechnic Institute for MALDI MS and synthetic peptides, and the fourth year honours students I've had the pleasure to influence: Ian Dobson, Paul Baba, Bao Do, Clement Poon and Sebastien Roussel.

Thanks from my heart to my ex-fiancée Sue Efler who has given me love and laughs for some of the best times and support through some of my worst.

To my brothers, Paul, James and Alex and finally I am ever grateful to my parents for their support through the years. They have made me who I am today.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

2D	two dimensional
αCHCA	α-cyano-4-hydroxycinnamic acid
αCT	α-chymotrypsin
μCi	microCurie
AGE	advanced glycation end-product
Ala	L-alanine
Arg	L-arginine
Asn	L-asparagine
Asp	L-aspartic acid
ATEE	*N*-acetyl-L-tyrosine ethyl ester
BME	*β*-mercaptoethanol
BSA	bovine serum albumin
CPMAS	cross polarization magic angle spinning
Cys	L-cysteine
D_2O	deuterium oxide
DD	dipolar dephasing
ddH_2O	distilled deionized water
DEPT	distortionless enhancement by polarization transfer
dH_2O	distilled water
DNA	deoxyribonucleic acid
Dns	5-dimethylamino-1-naphthalenesulfonyl
ESI	electrospray ionization

FAB	fast atom bombardment
Gln	L-glutamine
HCℓ	Hydrochloric acid
hHb	human hemoglobin
HI	Hydriodic acid
His	L-histidine
HPLC	high pressure liquid chromatography
HTLE	Hunter thin layer electrophoresis
HVPE	high voltage paper electrophoresis
ICAT	isotope coded affinity tag
Ile	L-isoleucine
ITC	isothiocyanate
kHz	kilohertz
Leu	L-leucine
Lys	L-lysine
K_M	Michaelis constant
kV	kilovolts
LpH	lyophilized pH
m/z	mass to charge ratio
MALDI	matrix assisted laser desorption ionization
MAS	magic angle spinning
mCi	milliCurie
MeOH	methanol

Met	L-methionine
MHz	megahertz
MS	mass spectrometry
mu	mass unit
MWCO	molecular weight cut-off
m/z	mass to charge ratio
N-Ac-D-Trp	*N*-acetyl-D-tryptophan
N-Ac-L-Trp	*N*-acetyl-L-tryptophan
NMR	nuclear magnetic resonance
PAGE	polyacrylamide gel electrophoresis
PITC	phenyl isothiocyanate
pH	negative logarithm of the proton concentration
Phe	L-phenylalanine
pK_a	negative logarithm of the acid ionization constant
PMM	peptide mass mapping
ppm	parts per million
Pro	L-proline
PSD	post source decay
Pth	phenylthiohydantoin
RNase	ribonuclease
Ser	L-serine
s/n	signal to noise ratio
TCA	trichloroacetic acid

TEA	triethylamine
TFA	trifluoroacetic acid
T_g	glass transition temperature
Thr	L-threonine
TMA	tetramethylammonium
TOF	time of flight
TPCK	*N*-tosyl-L-phenylalanine chloromethyl ketone
Trp	L-tryptophan
Tyr	L-tyrosine
UV-Vis	ultraviolet visible
Val	L-valine

Introduction

A protein's function is defined by its structure. Many techniques exist for determining a protein's structure; nuclear magnetic resonance spectroscopy (NMR), X-ray crystallography, optical rotation and, most recently, mass spectrometry are just a few examples. However, most of these techniques rely on some prior knowledge of the proteins function. For proteins of unknown function, the deciphering of protein structure-function relationships can be initially investigated by means of protein chemical modification. Two strategies currently exist for the structural modification of proteins to probe protein structure-function relationships. One, site directed mutagenesis (Winter *et al.*, 1982; Zoller and Smith, 1982; 1983) is a technique where a point mutation in the deoxyribonucleic acid (DNA) sequence encoding for the protein is made. This leads to a substitution of an amino acid after translation and can give insight as to the role of this amino acid in the protein's function. Although this method is powerful and is currently widely used, it relies on prior knowledge of the protein's sequence, it is labour intensive and it is sometimes difficult to interpret causal relationships. For example, an apparently benign substitution can lead to a dramatic effect on the protein's function. The other strategy used to study protein structure-function relationships is the chemical modification of proteins by the use of chemical reagents. It was at first thought that site-directed mutagenesis would replace chemical modification but subsequent experience has shown that these approaches complement each other.

Chemical modification of proteins is a well established approach for studying structure-function relationships in proteins (Means and Feeney, 1971; Darbre, 1986; Lundblad, 1995). Most reagents employed for chemical modification react with the proteins' nucleophilic functional groups and for a reagent to be successful in providing useful information on protein structure-function relationships it must possess specificity for one type of protein functional group over all others. The specificity of the reaction, to a limited extent, can be controlled by managing certain parameters such as; the time of reaction, the pH of the solution (*e.g.* targeting N-terminal α-amino groups (pKa ~8-9) vs. ε-amino groups (pKa ~10.5) by keeping the pH between 8.5 and 9) and the concentration of reagent(s). Another limiting factor is that in order to probe a protein's native structure, chemical modifications are usually carried out under aqueous conditions. Therefore the reagent and intermediates, if any, and the products formed must be reasonably stable in this aqueous environment. The reagent must also be reasonably miscible or soluble with water as a biphasic reaction would be slow and may lead to protein denaturation. In most cases water acts as a competing nucleophile and as a result reagents must be present in large excess for a reasonable extent of modification to occur. Consequently, the excess reagent must be quenched, the by-products removed and the modified protein purified either by dialysis, chromatography or other means. Other limitations imposed in performing the modification under aqueous conditions are the possibility of proteolytic damage to the protein of interest or other chemical modifications which may occur under the usual neutral or moderately alkaline aqueous conditions.

Lyophilization of proteins and peptides

Lyophilization (freeze drying) is a process where the solvent, usually water, is removed by sublimation which is accomplished by placing the frozen solution under vacuum. Many products can be lyophilized such as pharmaceutical drugs, enzymes for use in research and industry, synthetic peptides for pharmaceutical applications and protein, peptide or other formulations for drug delivery (Cleland *et al.*, 1994). Under the right conditions and using the appropriate co-lyophilized components, *i.e.* excipients, lyophilization of proteins and peptides can result in a considerably longer shelf-life versus being in solution. Lyophilization also eliminates the need for refrigeration and reduces other special handling procedures which ultimately lower costs.

The common excipients used in lyophilization are poly-hydroxy compounds, carbohydrates such as sucrose and other non-reducing sugars (Johnson *et al.*, 2002). These are thought to either replace the water during sublimation, the water replacement theory, or provide a glassy-like amorphous phase preventing phase transitions, the vitrification hypothesis. According to the water replacement theory, the hydrogen-bonding provided by the water that stabilizes the protein in solution is replaced by hydrogen-bonding with the poly-hydroxy excipients. During lyophilization, all solutes are freeze concentrated and as a consequence crystallization of buffers or other salts may occur which would induce irreversible change in the protein structure resulting in decreased stability. However, if the protein is maintained in an amorphous state using excipients, no phase transition occurs, therefore the protein structure is maintained. Both water replacement and preserving an amorphous-like material are thought to be involved in maintaining the stability of lyophilized proteins and peptides. The vitrificaton hypothesis is substantiated by the fact that lyophilized products have a glass transition temperature (T_g); therefore they are in a glass-like state. If the temperature of the lyophilizate increases above its T_g, the stability of the protein or peptide deteriorates. Different excipients at different mole ratios will result in different T_g values for the specific protein or peptide, for example trehalose provides a higher T_g than sorbitol (Souillac *et al.*, 2002). A humid atmosphere will also lower the T_g, causing instability of the lyophilized product (Bell *et al.*, 1995; Mazzobre *et al.*, 1997; Byrn *et al.*, 2001). Proteins themselves can serve as excipients and a common practice is to add a relatively inert protein to a preparation, such as albumin, to maintain the integrity of the lyophilized product.

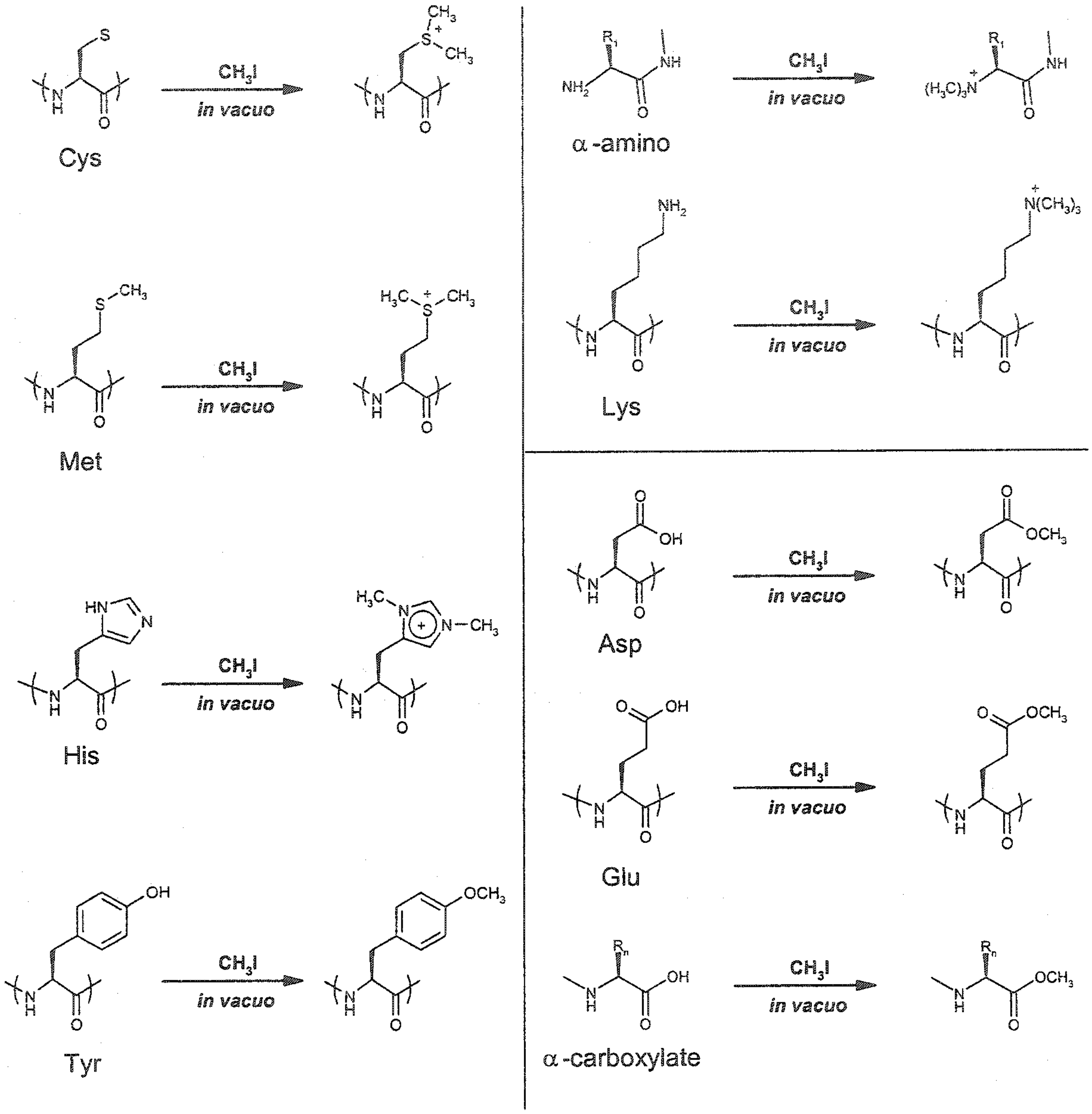
Without the use of excipients the lyophilization process in certain circumstances can itself lead to deterioration of the resulting activity. This can occur during one of three stages. First the initial freezing of the protein or peptide solution can cause irreversible damage due to phase separation of salts, buffers and certain excipients during ice formation (Randolph, 1997). Second the drying, a two stage process in which initial drying of the frozen solution by sublimation occurs followed by drying to ambient temperature under vacuum. Third, not part of the lyophilization process itself but important nevertheless is the re-solubilization step. Excipients can serve as "cryoprotectants" during the freezing step and "lyoprotectants" during drying. For re-solubilization the proper choice of buffers will re-establish the pH of the product solution which may have undergone some alteration during lyophilization. Although much progress has been made in understanding the processes involved in lyophilization and how different excipients in peptide and protein formulations play a role in stabilization, the addition of such additives, their selection and amounts is mostly chosen empirically.

It has been shown that lyophilized (freeze-dried) enzymes retain their catalytic activity when placed in organic solvents (Zaks and Klibanov, 1988; Klibanov, 1989; Koskinen and Klibanov, 1996). It can be inferred from this observation that proteins retain their structural integrity under nonaqueous conditions. The activity of the lyophilized protein, and therefore its structure, is dependent on the conditions prior to lyophilization. The pH of the solution from which the protein was lyophilized, termed "pH memory" (Zaks and Klibanov, 1988; Klibanov, 1995; Xu and Klibanov, 1996; Klibanov, 1997) and the use of co-lyophilized additives (excipients) (Russell and Klibanov, 1988; Braco *et al.*, 1990) both play an important role in the resulting biological activity. For example, the pH optima and catalytic activity of two serine proteases; subtilisin and αCT, in a non-aqueous organic solvent are dependent upon the pH of lyophilization and correlates with their aqueous enzymatic pH optima (Xu and Klibanov, 1996) Furthermore, lyophilized proteins in organic solvents have an enhanced thermo-stability, a property allowing for possible elevation in temperature to accelerate the catalytic reaction without disrupting the native conformation of proteins (Zaks and Klibanov, 1988; Broos *et al.*, 1995).

In vacuo chemical modification of proteins

Chemical modification of proteins to investigate structure-function relationships (Means and Feeney, 1971; Darbre, 1986; Lundblad, 1995) is a powerful technique but it is inherently limited due partly to the fact of having to perform the reaction under aqueous conditions to probe native protein structure. Although protein chemical modification can be performed with the addition of a miscible organic solvent, to aid in the solubility of the reagent, there is a limit before a change in protein structure occurs with possible protein denaturation. Consequently, this places limits in the choice of chemical reagent and, as a result, in the types of modifications possible. To circumvent this disadvantage, Taralp and Kaplan devised a method in which proteins are chemically modified in the absence of any solvent (Kaplan and Taralp, 1997; Taralp and Kaplan, 1997). The lyophilized protein is modified with a gaseous reagent under vacuum (see Figure 0-1). In performing the reaction under non-aqueous conditions many problems inherent to aqueous modifications such as, proteolysis, protein conformational dynamics, and possible protein denaturation are avoided. Furthermore, this method allows for the use of reagents which would otherwise not be compatible with water and for the presence of possible intermediates formed not stable in the presence of water. The thermo-stability of lyophilized proteins allows for the use of higher temperatures to increase the rate of reaction and to keep the reagent in the gas phase. Taking advantage of the "pH memory" effect of lyophilized proteins *i.e.* the ionizable functional groups have the same degree of ionization as in the aqueous solution but there is no exchange of protons, the extent of modification can be controlled simply by adjusting the pH of the solution prior to lyophilization (LpH). In the method, only a small amount of reagent is required and it can be easily removed and recovered allowing for the cost effective use of isotopically enriched reagents to aid in characterization. For example, Taralp and Kaplan used ^{13}C-iodomethane to investigate the *in vacuo* reaction of this reagent with lyophilized proteins by ^{13}C NMR spectroscopy. The modifications afforded with the *in vacuo* methylation of proteins with iodomethane is shown in Scheme 0-1. Methylation of ionizable functional groups can be controlled by the LpH. The sulphur containing amino acids; cysteine and methionines, are converted to their dimethylsulfonium derivatives. The N-terminal α-amino group and the ε-amino group of lysine react with iodomethane to form predominantly the trimethylammonium derivatives, in a pH dependent fashion.

The imidazole containing amino acid; histidine, is converted to its dimethylimidazolium derivative, the phenolic residue of tyrosine, at higher LpH with protein unfolding is methylated to form a methyl ether. Carboxylate functional groups of the C-terminal α-carboxyl, the β-carboxyl group of aspartic acid and the γ-carboxyl group of glutamic acid are converted to their respective methyl esters. In the initial publications, the esterification of carboxylate functional groups was not observed in the NMR spectra due perhaps to the overlap of the strong chemical shift signals of the trimethylammonium derivatives of amino groups and the possible hydrolysis of the methyl esters prior to NMR analysis. The extent of esterification was later investigated and compared to a novel *in vacuo* esterification method using methanol HCℓ (Vakos *et al.*, 2001).



Scheme 0-1 Predominant derivatives formed by the *in vacuo* methylation of proteins with iodomethane

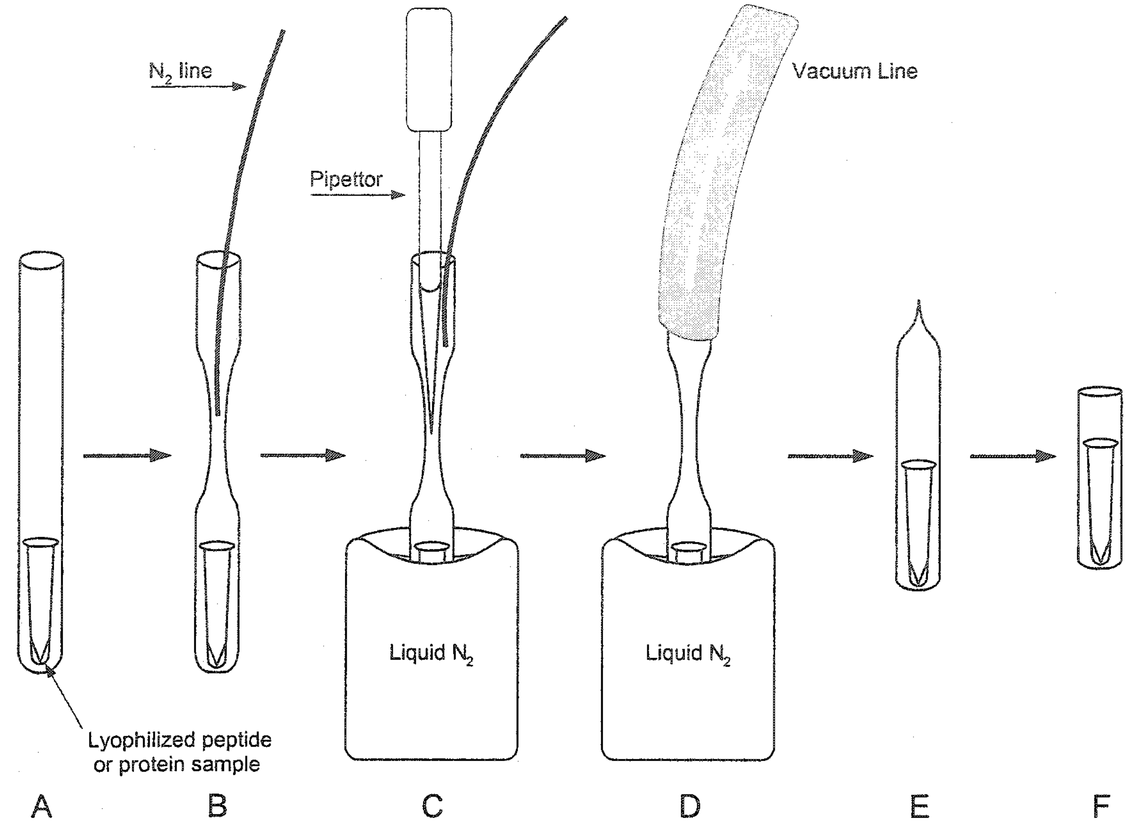


Figure 0-1 Schematic diagram of the *in vacuo* methylation method.

A: Lyophilized peptide or protein sample in a small glass or plastic insert is placed into larger borosilicate glass tube, B: Outer tube is narrowed with a flame and purged with nitrogen, C: Iodomethane is introduced under a nitrogen atmosphere with a pipettor by immersing the vessel in liquid nitrogen, D: Reaction vessel is placed under vacuum while immersed in liquid nitrogen, and E: sealed with a flame and placed in a block heater for reaction to proceed. F: Reaction is terminated by trapping out excess reagent and breaking the vacuum.

Chapter 1 Increased detection of peptides in MALDI MS by trimethylation of their amino groups

Introduction

MALDI MS

A brief description of matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS) is presented here as many excellent reviews have recently been published on its use in proteomics (Godovac-Zimmermann and Brown, 2001; Mann *et al.*, 2001), protein chemistry (Roepstorff, 2000) and other biological applications (Matsuo and Seyama, 2000). MALDI was developed in the middle to late 1980s (Karas and Hillenkamp, 1988) with the concurrent development of electrospray ionization (ESI) (Fenn *et al.*, 1989) as "soft ionization" methods to determine the mass of intact proteins by MS. Both methods allow for the sensitive determination of proteins and peptides, but ESI MS is more susceptible to the ion suppression effect of any contaminating salts and therefore usually requires a time consuming chromatographic separation procedure prior to MS. MALDI MS in its nature is less prone to the effect of salts and as a result samples can be first applied to MALDI in a high throughput manner prior to further investigation by ESI.

The sample (analyte) is co-crystallized with an excess of a UV (or IR) absorbing matrix and allowed to dry on a metal sample plate (probe tip) (see Figure 1-1). The choice of matrix depends on the type of analyte and on its size (peptide *vs*. protein). Gas phase charged molecules of the analyte are produced by the irradiation of the matrix with a pulse of a laser which absorbs the laser energy and "transfers" it to the analyte which in turn is ionized and "flies" into the vacuum. The ions are accelerated through a strong electric field giving them all the same kinetic energy. Depending on their mass to charge ratio, they have different velocities and thus different time of flight (TOF) after traversing a fixed field-free distance. The resulting TOF spectrum is calibrated with standards preferably within the sample (internal calibration) or as a separate sample (external calibration).

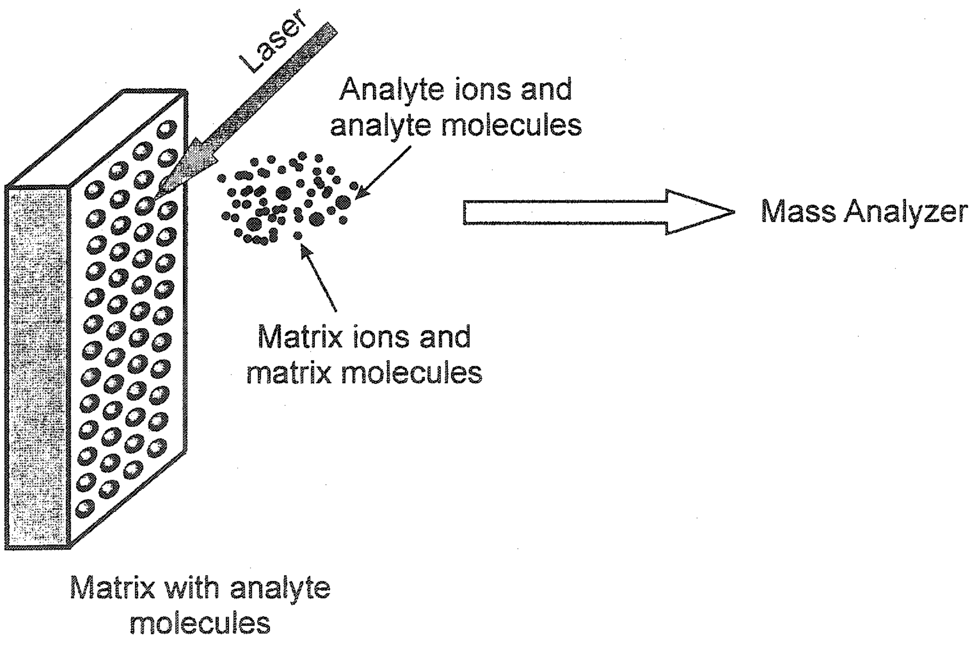


Figure 1-1 Schematic of MALDI sample plate

The recent elucidation of the human (Lander *et al.*, 2001) and other genomes (www.ncbi.nlm.nih.gov), has spurred efforts to identify all the proteins expressed by cells in various tissues, *i.e.* the proteome. Typically, proteins are separated by 2D polyacrylamide gel electrophoresis (2D-PAGE), the proteins of interest are cut out of the gel, enzymatically digested (Katayama *et al.*, 2001) and the proteolytic peptides produced are subjected to MS producing a peptide mass map (Cohen and Chait, 1997; Ehring *et al.*, 1997). The peptide mass map is searched for hits in a complementary DNA (cDNA) database and matched, with confidence limits, to a theoretical peptide mass map produced from the specific enzymatic cleavage of known encoded proteins.

MALDI-TOF MS is one of the major tools being employed in proteomics for the analysis of peptides generated by enzymatic digestion of expressed proteins. Trypsin is the most commonly used enzyme for this purpose since the predictability of its cleavage sites facilitates identification of proteins by database searching. MALDI MS is preferred over other soft ionization MS methods (*i.e.* ESI) as small amounts of salts, detergents, and other contaminants are tolerated (Bornsen, 2000) and simpler spectra are generated due to the fact that singly charged species are observed (Karas *et al.*, 2000). One of the difficulties encountered with this strategy, however, is that there is a great variability in the signal intensity for the different peptide fragments generated and non-peptide impurities can be misleading in database searching. In many cases, not all the predicted fragments are observed, reducing the chances of success and the reliability of identification using database searches. It has been observed that arginine-containing peptides are more easily detected than lysine-containing peptides (Krause *et al.*, 1999). Conversion of lysine residues in peptides to homoarginine by guanidinylation of ε-amino groups (see Scheme 1-1) has been shown to improve their detection to the same level as that of arginine-containing peptides (Beardsley *et al.*, 2000; Brancia *et al.*, 2000; Hale *et al.*, 2000). An even greater detection has been claimed with the reagent 2-methoxy-4,5-dihydro-1*H*-imidazole (Peters *et al.*, 2001) (see Scheme 1-1B). It is believed that the higher pK_a of the guanidino group (pK_a ca. 12) as compared to that of the ε-amino amino (pK_a ca. 10.5) (Creighton, 1993) contributes to the stability of charged mono-protonated peptide ion species in the gas phase leading to greater signal intensity.

Other methods for improving peptide detection have mainly focused on sample application (Landry *et al.*, 2000) and experimenting with different matrix compositions (Billeci and Stults, 1993; Gusev *et al.*, 1995; Bornsen, 2000).

Isotopic labelling as a means of simplifying spectral interpretation

Isotopic labelling has been used as a means of simplifying spectral interpretation. For example, trypsin digestion in a mixture of $H_2{}^{16}O$ and $H_2{}^{18}O$ has been attempted to produce pairs of tryptic peptides differing in their mass by two mass units (Stewart *et al.*, 2001). However, this method found limited success due to the fact that the peptides produced are themselves substrates for trypsin and as such undergo additional nucleophilic attack by water at the C-terminal carboxylate resulting in a mass distribution (Gevaert *et al.*, 1997; Shevchenko *et al.*, 1997; Stewart *et al.*, 2001) and not a well defined doublet. Another example of isotopic labelling to simplify mass spectra is the introduction of an N-terminal label with the "fragmentation-directing" moiety nicotinyl-*N*-hydroxysuccinimide for MS/MS. In MS/MS peptide ions to be sequenced are chosen in one mass spectrometer and fragmented by the collision of a gas and the fragments generated are sent to a second mass spectrometer. Fragmentation of the peptide backbone occurs (see Figure 1-2) at three possible sites.

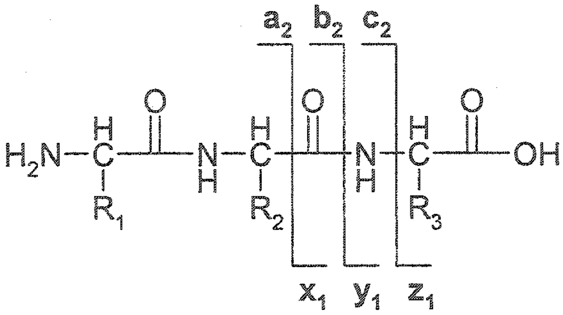


Figure 1-2 Peptide backbone fragmentation nomenclature in MS/MS (taken from (Chapman, 2000)).
Fragmentations producing N-terminal peptides are denoted a, b and c with subscript numbering starting from the N-terminal and the same fragmentations producing C-terminal peptides are denoted x, y and z with subscript numbering starting from the C-terminal end.

An isotopic mixture of H_4/D_4 labelled reagent was used to increase the relative yield of the b-ion series during MS/MS. It is interesting to note that this increase in b-ion formation was attributed to the "highly basic nature" of the nicotinyl group yet no effect on MALDI MS detection was mentioned (Münchbach *et al.*, 2000). Isotope coded affinity tag (ICAT) reagents have also been used to label free sulfhydryl groups of cysteinyl residues in protein mixtures (Gygi *et al.*, 1999; Gygi *et al.*, 2000; Griffin *et al.*, 2001).

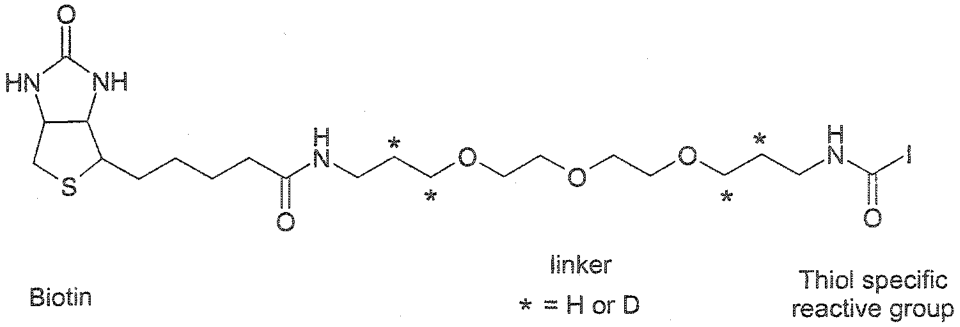
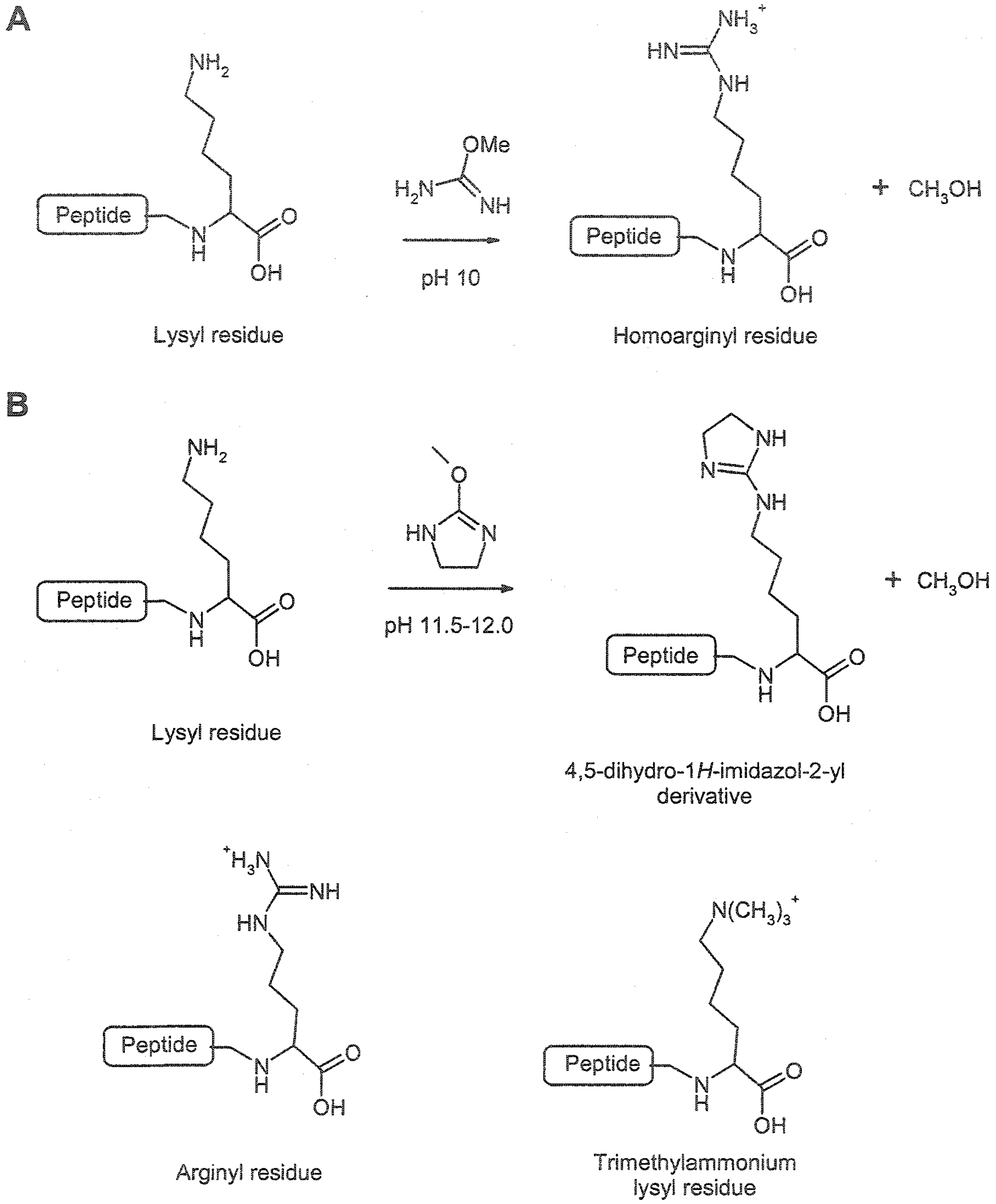


Figure 1-3 Structure of the ICAT reagent

The reagent consists of three elements: a thiol specific reactive group, a linker which can incorporate stable isotopes and the biotin moiety (see Figure 1-3) which allows for the isolation of the labelled peptides, produced by an enzymatic digest, by the strong affinity for the binding of biotin by avidin immobilized onto a column. This technique has shown to be effective but relies on many separation procedures and its greatest limitation is that it relies on the existence of free sulfhydryl groups of cysteinyl residues to produce labelled peptides. These groups are usually oxidized to various degrees or poorly accessible; buried within the protein structure or disulfide linked, and therefore must be modified under reduced denaturing conditions. Also the occurrence of cysteine residues in proteins is itself limited. The methods mentioned above all suffer from the same major disadvantage in that they are performed under aqueous conditions. In the case of peptides this may not allow for complete derivatization as some peptides are poorly or not soluble in aqueous media. The reagent must be present in excess as water usually competes during the reaction and therefore a buffer, inert to the reagent is required to maintain the pH. As a result, excess reagent, by-products and salts/buffer must be removed by reverse phase chromatography prior to MS. Peptide mass mapping (PMM) mass tagging of proteins with stable isotope-labelled amino acids during MS can be used to identify peptides from the map. Labelled and un-labelled amino acid precursors are mixed in different ratios and added to the growth medium (Hunter *et al.*, 2001). Although this is an attractive strategy it is costly, labour intensive and cannot be applied for true unknowns as a pre-existing *in vitro* model system is required.

Methylation of amino groups in proteins and peptides can be easily achieved by the *in vacuo* reaction with iodomethane (Kaplan and Taralp, 1997; Taralp and Kaplan, 1997; Vakos *et al.*, 2000; Vakos *et al.*, 2001). This reaction produces a trimethylated quaternary ammonium derivative with a permanent positive charge (see Scheme 0-1). If the increased sensitivity for the MALDI MS detection of peptides containing guanidino groups is due to the ability of this group to retain a positive charge, then conversion of the peptide's amino group to the trimethylammonium derivative should also show a similar effect. *In vacuo* reaction with iodomethane requires only one step with no solvent removal after completion of the reaction. Therefore, the methylation of very small amounts of peptide or protein can be carried out without loss of sample. This chapter describes the effects of trimethylation of amino groups on the detection of peptides by MALDI MS and demonstrates how this derivatization can be used with a mixture of isotopes to differentiate between peptide and non-peptide signals in the mass spectrum.



Scheme 1-1 **A** Conversion of lysyl residues to homoarginine with *O*-methylisourea and **B** Reaction of lysyl residue with 2-methoxy-4,5-dihydro-1*H*-imidazole. The structures of the arginyl residue and the trimethylammonium derivative of the lysyl residue are also shown for comparison

Materials and Methods

The following peptides: Angiotensin II Fragment 3-8: H-Val-Tyr-Ile-His-Pro-Phe-OH (98% pure, Lot#106H5870), H-Arg-Gly-Glu-Ser-OH (99% pure, Lot#78H5806), H-Arg-Gly-Pro-Phe-Pro-Ile-OH (97% pure, Lot#115H0838), H-Gly-Phe-Asp-Leu-Asn-Gly-Gly-Gly-Val-Gly-OH (99% pure, Lot#30H0099), Laminin Fragment 929-933: H-Tyr-Ile-Gly-Ser-Arg-OH (99% pure, Lot#50K1570), Leucine Enkephalin: H-Tyr-Gly-Gly-Phe-Leu-OH (Lot#10K5925), oxidized B-chain from bovine insulin (insulin B_{ox}-chain) (81% pure, Lot#38H5805), H-Phe-Leu-Glu-Glu-Leu-OH (99% pure, Lot#88H5821), and H-Val-Gly-Val-Ala-Pro-Gly-OH (99% pure, Lot#103H58754) were obtained from Sigma and H-Ser-Gln-Asn-Tyr-Pro-Ile-Val-OH (>99% pure (HPLC) acetate salt, Lot#524741) was obtained from Bachem AG. All were used without further purification.

The following proteins: bovine serum albumin (BSA) (Lot#78C0388), human hemoglobin (hHb) (Lot#67H9306), papain (from papaya latex, suspension in 0.05 M sodium acetate, 24 units/mg protein, Lot#124F8045) and trypsin from porcine pancreas (Lot#53H07111) were obtained from Sigma.

The following chemicals: *N*-tosyl-L-phenylalanine chloromethyl ketone (TPCK), tetramethylammonium (TMA) phosphate pentahydrate, triethylamine (TEA) (99%), boric acid, citric acid, hydriodic acid (HI), sulphuric acid and trifluoroacetic acid (TFA) were obtained from Sigma. Trichloroacetic acid (TCA) and hydrochloric acid (HCℓ) were acquired from Aristar. Iodomethane was purchased from BDH, iodomethane-^{13}C (99 atom%) and iodomethane-d_3 (99+ atom%) were obtained from Aldrich. All other chemicals employed were of reagent grade. pH adjustments were made using a

Radiometer Copenhagen type PHM-26 pH meter fitted with an Orion semi-micro gel-filled combination electrode model 91015.

Methylation of peptides by the *in vacuo* method

In vacuo methylation of peptides with iodomethane was carried out similarly to the procedure previously described in its application on lyophilized proteins (Kaplan and Taralp, 1997; Taralp and Kaplan, 1997) with modifications described below (see Figure 0-1). Peptide samples (typically 0.1 mg), excluding insulin B_{ox}-chain and those lyophilized under acidic conditions (*vide supra*), were dissolved in 100 µL of 10 mM TMA phosphate buffer (pH 8.0, 9.0 or 10.0). A stock solution of the buffer was prepared by dissolving the pentahydrate of TMA hydroxide for a final concentration of 10 mM and adjusting the pH with 100 mM phosphoric acid before final volume adjustment. The samples were aliquoted and lyophilized from small borosilicate glass conical vials (Kimble®, 13 x 45 mm) which were then placed in larger glass tubes (15 x 100 mm). After forming a narrow constriction near the upper portion of the outer tube with a flame, the lower portion of the tube was immersed in liquid nitrogen and iodomethane (30 µL) was introduced under a dry nitrogen atmosphere. The tube was evacuated (ca. 50-70 mTorr) and sealed under vacuum at the constriction. The sealed tubes were placed in a mineral oil block heater at 75°C for 18 h. The reaction was stopped by trapping out the excess reagent at the top of the tube by fitting a rubber jacket (a bored out rubber septa) over the top of the tube and filling it with liquid nitrogen. The vacuum was released and the tube opened by simply scoring the glass under the jacket and touching the score mark with the white-hot tip of a glass rod. The inner glass tube containing the methylated lyophilized peptide was easily removed with a pair of tweezers and placed into micro-centrifuge tubes for the ease of handling. Non-methylated peptide samples serving as controls were handled in a similar fashion without the introduction of iodomethane.

Removal of methyl esters

After the *in vacuo* methylation procedure described above, some samples were dissolved in 100 µL of 2% (v/v) TEA, transferred to micro-centrifuge tubes, heated in a water bath for 1 hour at 90°C and dried under vacuum using a Speed-Vac™ (Savant) centrifuge apparatus.

Methylation of insulin B_{ox}-chain

Three separate samples each containing insulin B_{ox}-chain (0.86 mg, ca. 200 nmol) in 100 µL 2% (v/v) TEA were lyophilized in small glass tubes. *In vacuo* methylation was performed with CH_3I and $^{13}CH_3I$ as described above using the tubes with the lyophilized samples as inserts. As a control, the B_{ox}-chain was heated *in vacuo* without iodomethane.

After *in vacuo* methylation, methyl esters were removed by the procedure described above. Samples were left to cool to room temperature. Aliquots (50 µL) were transferred to micro-centrifuge tubes and neutralized with 5% (v/v) formic acid (5 µL). Mixtures containing equal amounts of non-methylated and ^{12}C-methylated B_{ox}-chain and, ^{12}C and ^{13}C methylated B_{ox}-chain, were also prepared.

Trypsin digestion

TPCK trypsin was prepared as previously described (Schoellmann and Shaw, 1963). A stock solution of TPCK trypsin was prepared as follows: $CaCl_2$ (5 μL of 1M) was added to NH_4HCO_3 (2.0 mL of 200 mM) and TPCK trypsin (50 μL of 5 mg/mL) was added. An aliquot (100 μL) of this stock was added to each of the B_{ox}-chain samples (ca. 50:1 w/w; B_{ox}-chain to enzyme); digestion was carried out for 1h at 37°C in a shaker bath. Samples were dried using a Speed-Vac™ centrifuge apparatus. They were then dissolved and dried from ddH_2O (20 μL) and then re-dissolved and dried from 5% formic acid (10 μL) to acidify prior to MS.

Methylation of proteins

BSA, 500 mg and hHb, 100 mg were each dissolved in ddH_2O (10 mL) and dialyzed (MWCO 3500) overnight at 4°C against dH_2O (5 x 4L). After dialysis both were made up to a final volume of 50 mL with ddH_2O. From these stocks two sets of solutions (1mg/mL, 10 mL total) were prepared and the pH adjusted to 7.5 and 8.0 with 1M NaOH *via* plastic micro-syringe. Protein solution was transferred to borosilicate hydrolysis tubes in 1mL aliquots, frozen in liquid nitrogen and lyophilized overnight.

In vacuo methylation was performed with CH_3I and CD_3I (30 μL) as described (*vide infra*). After reaction the methylated protein samples and heated blanks were dissolved in 5% formic acid (100 μL) and transferred with one wash (5% formic acid 100 μL) to micro-centrifuge tubes.

Hydrolysis of methyl esters

The removal of methyl esters was performed similarly to above but all samples were dissolved in 2% pyridine (200 μL) and heated in a water bath for 30 min at 90°C. Mixtures containing equal amounts of non-methylated and CH_3 methylated protein, and CH_3 and CD_3 methylated protein were also prepared. Samples were then dried under vacuum using a Speed-Vac™ centrifuge apparatus.

Papain digest

A papain stock solution was prepared (1 mg/mL in pH 6.5 volatile buffer (glacial acetic acid/pyridine/dH_2O (3:100:900) with 1% (v/v) BME). Protein samples were dissolved in pH 6.5 volatile buffer (500 μL) and to these an aliquot (20 μL) of the papain stock solution was added. Digestion was carried out for up to 4 h at 37°C in a shaker bath. Aliquots (100 μL) were taken at 15 min, 30 min, 1 h, 2 h, and 4 h and the digestion stopped by adding them to 5% formic acid (100 μL) and immediately freezing them in liquid nitrogen. Once all time points were collected samples were dried by Speed-Vac™.

Methylation of a peptide lyophilized from acids

H-Ser-Gln-Asn-Tyr-Pro-Ile-Val-OH obtained as the TFA salt was dissolved in ddH_2O (1 nmol/50 μL) and aliquoted (50 μL) to 1.5 mL polyethylene micro-centrifuge tubes and lyophilized. The following acids: TCA, TFA, HCℓ, HI, acetic, formic, boric, citric and sulfuric were prepared in 13 mM stocks (concentration equivalent to 0.1% (v/v) TFA). Lyophilized peptide samples were re-dissolved and lyophilized twice for each acid. *In vacuo* methylation was performed as previously described by placing the micro-centrifuge tube (cap removed) directly into the glass tube to be narrowed. Samples were methylated for 48 hrs at 100°C. An aliquot of the methylated TFA lyophilized peptide was mixed to an equal aliquot of CD_3 methylated (24 hrs, 75°C) pH 8.0 (10 mM TMA phosphate) lyophilized peptide.

MALDI MS

Samples were dissolved in an appropriate volume of 5% formic acid (10–100 μL) and dried under vacuum using a Speed-Vac apparatus. Samples were prepared for MALDI-TOF MS by dissolving in 0.1% TFA and mixing with the matrix, α-cyano-4-hydroxycinnamic acid (α-CHCA) (3 mg dissolved in 300 μL of acetonitrile ethanol (1:1)). The spectra in Figure 1-20, Figure 1-21 and Figure 1-22 were obtained using a PerSeptive Biosystems Voyager-DE™ Pro Biospectrometry™ Workstation MALDI-TOF MS; all other spectra were obtained on a Micromass TofSpec-2E™ MALDI-TOF MS. Prediction of peptide mass was done with the aid of the PAWS™ program (version 2001.08.14), Proteometrics, Inc.

Results and Discussion

Choice of lyophilization buffer

In the *in vacuo* method the pH of the lyophilized protein, the "LpH", is controlled simply by adjusting the pH of the protein solution prior to lyophilization with NaOH or HCℓ. Controlling the pH in this manner provides a simple means of controlling the extent of reaction of certain functional groups mainly α-amino vs. ε-amino groups.

However, for peptides, the amounts used (sub-nmol) and the desire to make this a method for general applications (*i.e.* proteomics) where peptide amounts may not be known and certainly at the pmol level or lower, the pH was controlled by the addition of a buffer. The choice of buffer is somewhat arbitrary and has not been fully investigated. It should however be alkaline to target the peptides' amino groups and perhaps volatile to limit the amount of remaining salt once lyophilized which may interfere during subsequent MS. A volatile buffer, at first, seemed to be the logical choice as presumably only the counter-ions to the peptide would remain however, after lyophilization the final LpH of the peptide-buffer mixture cannot be predicted *e.g.* a pH drop may occur (Zacharis *et al.*, 1999). Moreover, when TEA (2%) was tried some peptide-buffer lyophilizates were hygroscopic further complicating matters.

TMA-phosphate (10 mM) was therefore arbitrarily chosen as a non-volatile, non-hygroscopic buffer, which should not undergo a pH drop as a volatile buffer might or as the sodium phosphate buffer would (Randolph, 1997). Buffers with sodium or potassium were avoided as these tend to form positively charged adducts in MS potentially complicating the interpretation of peptide mass spectra. A non-volatile buffer was chosen for the simple reason that the lyophilized product could be observed as a dry, white residue while in many instances a volatile buffer left no observable residue with the question of sample loss. The TMA cation is unreactive towards iodomethane and therefore should not compete for the reagent during methylation whereas a volatile tertiary amine might react which would result in the alteration of the LpH and perhaps result in a limited yield of methylated peptide. Furthermore, TMA hydroxide, when heated at a higher temperature under vacuum, sublimes to trimethylamine, methanol and water (O'Neil and Merck & Co., 2001). Therefore excess TMA could have been potentially removed simply by heating *in vacuo* if the resulting salt proved to be a problem in MS. The phosphate counter ion does react with iodomethane to a certain extent (Vakos, 2001) as most anions would and this reaction cannot be avoided. However, the resulting product is neutral and expected not to interfere with subsequent MS. TEA was the first choice in solubilizing peptides as all peptides and proteins tested were completely soluble in TEA. The use of TEA in solubilizing proteins is especially attractive in that most proteins are completely denatured under highly alkaline conditions providing iodomethane access to potentially inaccessible (buried) functional groups. Highly alkaline conditions would not be prudent in an aqueous environment as prolonged exposure could lead to possible protein/peptide degradative reactions. However, under nonaqueous conditions, the protein solution is quickly frozen and lyophilized to minimize any possible degradation and was not observed with the model peptides used. Although this buffer was not ideal for peptides, it seems promising for proteins and should be further investigated for *in vacuo* chemical modifications.

Effect of trimethylation of peptide α-amino groups on MALDI MS detection

To observe the effect of introducing a permanent positive charge on peptide detection by MALDI MS by trimethylation of the α-amino group, peptides with only one amino group and no other reactive side chains other than carboxylate functional groups were chosen. The side chain of lysine, histidine, methionine or cysteine would be methylated to a positively charged derivative. The effect of the introduction of this other positive charge might complicate the interpretation of the resulting MALDI mass spectrum. For each of the following spectra, peaks are labelled as their mono-isotopic mass and an accompanying peak, one mass unit higher, arising from the natural abundance ^{13}C (~1%) may also be labelled. Although distilled deionized water was used to prepare and re-solublize the peptide samples it is common to observe the sodium adduct, $[M + Na]^+$(+22 mu from $[M + H]^+$) and the potassium adduct, $[M + K]^+$ (+34 mu from $[M + H]^+$)of the peptide. Due to different ionization yields of different peptides, and other ions, as explained earlier, peak height is not related to the amount of that specific analyte in the sample. In fact, peak height can be misleading as a major component of a sample may not be observed if it does not ionize under MALDI (*e.g.* a neutral species or negatively charged ion) and conversely a small amount of a component in the sample which ionizes well will dominate the spectrum. Furthermore, different spectra (*i.e.* different samples/spots on the MALDI plate) cannot be compared to one another since different amounts of salts, buffers, different concentrations and other components will have an effect on the crystallization of the matrix. As the co-crystallization of the matrix with the analyte is essential for MALDI, any perturbation in that process will greatly influence the resulting spectrum. Therefore to compare the trimethylammonium derivatized peptide to the non-derivatized peptide, equal amounts of both were combined and spotted on the MALDI plate.

The MALDI MS of H-FLEEL-OH (5 pmol) alone is shown in Figure 1-4. This peptide gives a signal at 650.7 m/z for the $[M + H]^+$ ion with minor impurities at 630.5 and 692.4 m/z, the $[M + Na]^+$ adduct is not observed. The MALDI MS signal intensity for this peptide is poor as the noise is relatively high. This may be due to the fact that this peptide possesses three acidic carboxylate groups to one basic N-terminal α-amino group.

The reaction scheme for this peptide with iodomethane is shown in Scheme 1-2. *In vacuo* reaction of H-FLEEL-OH with iodomethane trimethylates the α-amino of phenylalanine and methylates the α-carboxyl of leucine and the γ-carboxyl groups of glutamic acid. The MALDI MS of an equal mixture of the resulting peptide $(CH_3)_3N^+$-FLE(OCH_3)E(OCH_3)L-OCH_3 and the underivatized peptide is shown in Figure 1-6.

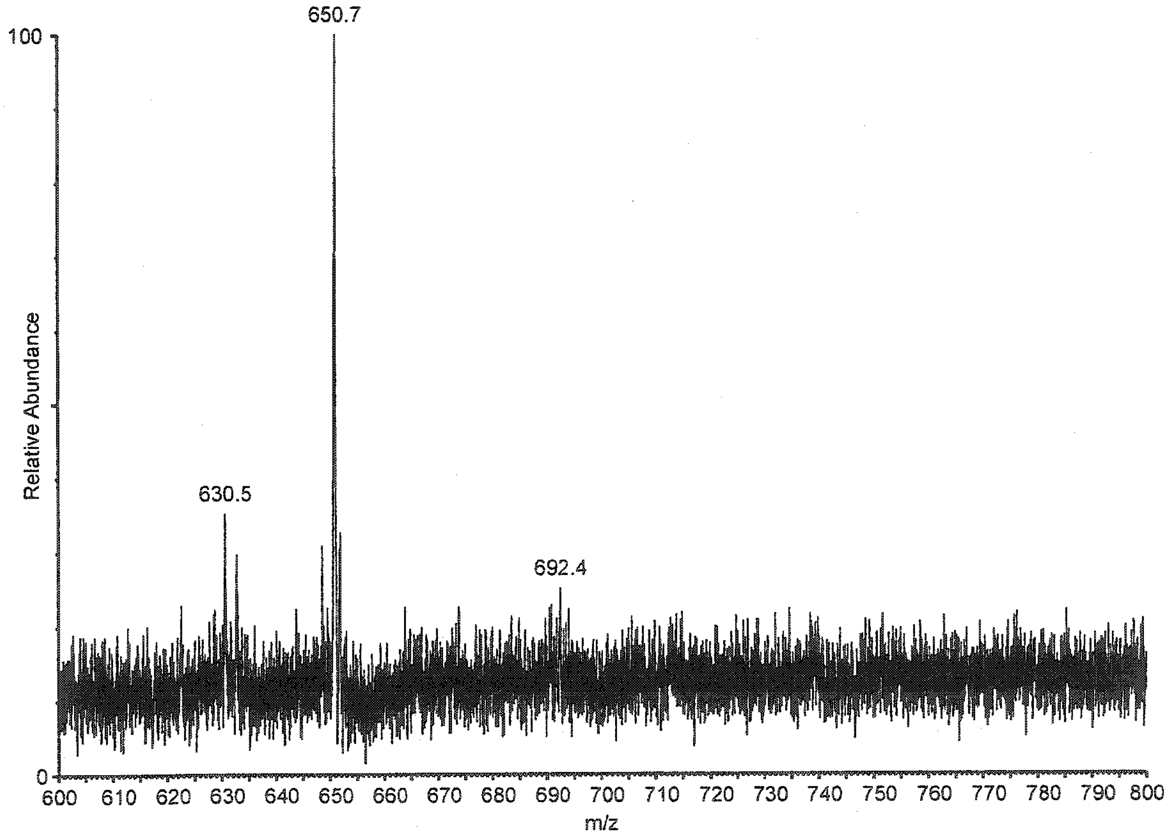
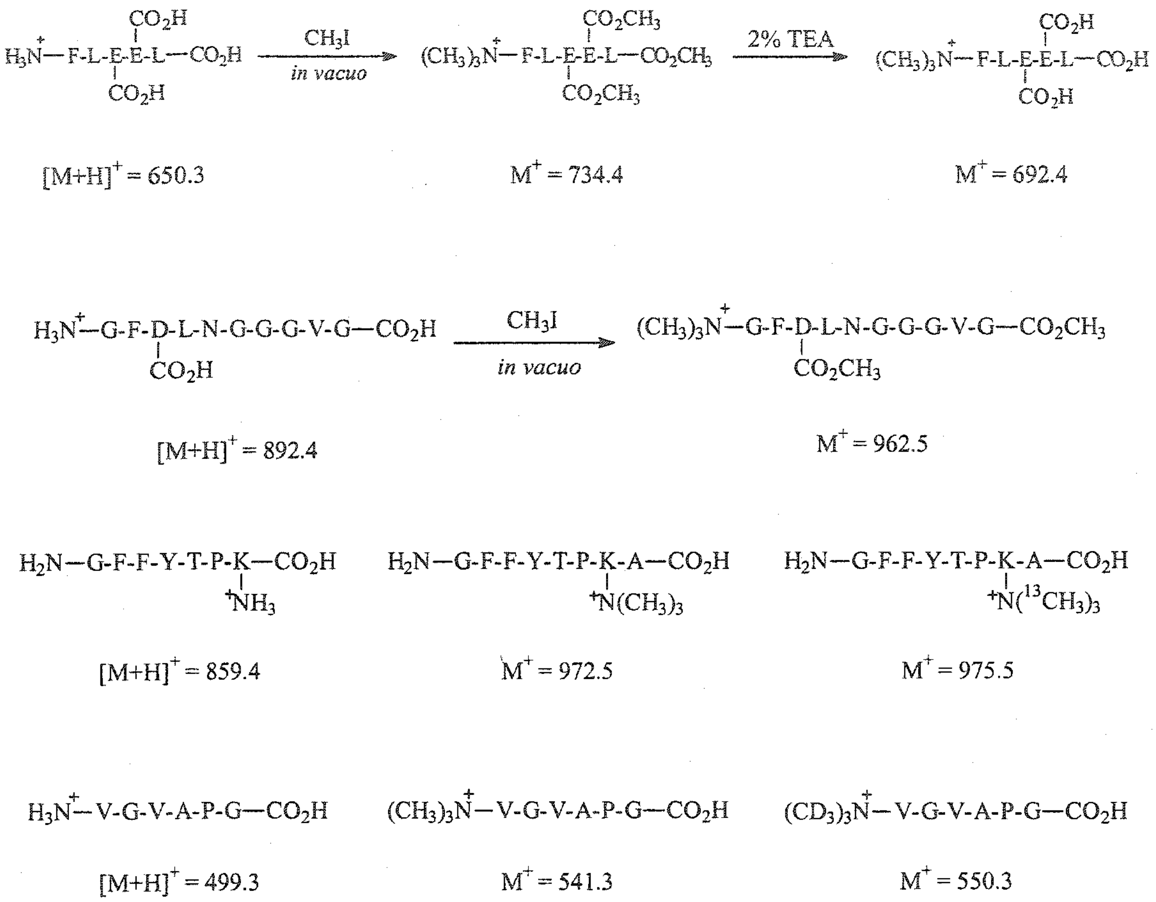


Figure 1-4 MALDI MS of H-FLEEL-OH $[M + H]^+ = 650.7$ m/z

In order to isolate the effect of trimethylation of the amino group on peptide detection, the methyl esters were removed by aqueous base hydrolysis to give only

$(CH_3)_3N^+$-FLEEL-OH; its MALDI MS spectrum is shown in Figure 1-5. The MALDI MS of an equal mixture of this peptide and the underivatized peptide is shown in Figure 1-7.



Scheme 1-2 Predicted mono-isotopic masses of test peptides and of *in vacuo* methylated test peptides

The effect of methylation of the α-amino group on the MALDI detection of H-FLEEL-OH is readily apparent. When an equal amount, *viz.* 10 pmols, of H-FLEEL-OH and $(CH_3)_3N^+$-FLEEL-OH are subjected to MALDI MS, the trimethylated α-amino derivative gives a much greater signal intensity than the underivatized peptide possessing a protonated ammonium group.

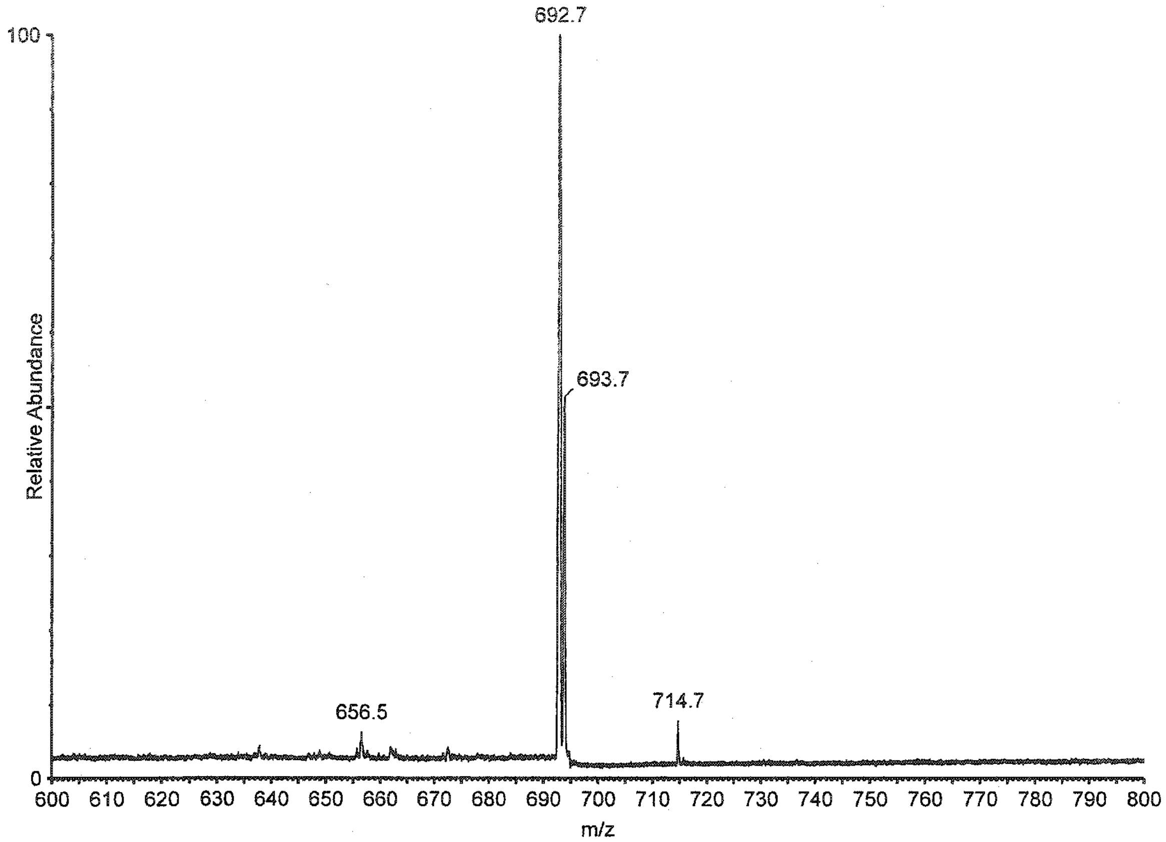


Figure 1-5 MALDI MS of $(CH_3)_3N^+$-FLEEL-OH $[M]^+ = 692.7$ m/z, $[M + Na]^+ = 714.7$ m/z

Comparison of the MALDI MS on equal amounts of the fully methylated derivative $(CH_3)_3N^+$-FLE(OCH_3)E(OCH_3)L-OCH_3 and the corresponding derivative in which the ester groups have been removed by base hydrolysis $(CH_3)_3N^+$-FLEEL-OH shows that the esterified derivative has approximately twice the signal intensity of the non-esterified derivative (Figure 1-8). Peaks at 707.0 and 721.0 m/z correspond to peptide derivatives containing four and five methyl groups, respectively. These peaks may arise either from the mono and di-ester derivative due to a small amount of hydrolysis during sample preparation prior to MS or an incomplete methylation of the amino group (*i.e.* mono and dimethyl amino) with methyl esters would result in derivatives with the same m/z. However, when the esters have been removed only the trimethylated derivative at 692.7

m/z and its corresponding sodium adduct at 714.7 m/z are observed (see Figure 1-5) confirming that these peaks are indeed due to incomplete esterification. While esterification of carboxyl groups does increase the signal intensity, it is clear that the major increase is due to the trimethylammonium group. It is not necessary to remove methyl esters but their removal by hydrolysis is easy and will simplify interpretation of spectra in cases where esterification is incomplete or when double isotope labelling is used (*vide infra*). It is also interesting to note that the sodium adduct is observed even though the presence of sodium was avoided.

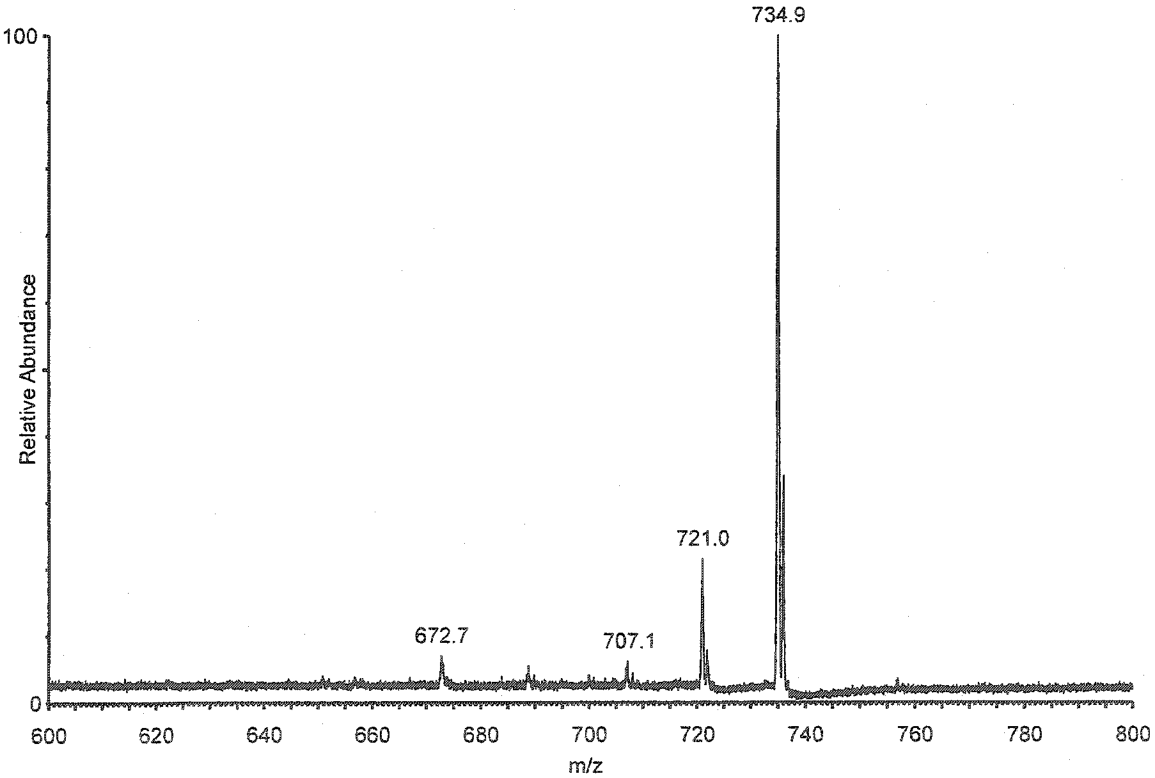


Figure 1-6 MALDI MS of an equal mixture of H-FLEEL-OH $[M + Na]^+ = 672.7$ m/z and $(CH_3)_3N^+$-FLE(OCH_3)E(OCH_3)L-OCH_3 $[M]^+ = 734.9$ m/z

Partial hydrolysis of one methyl ester (-14 mu) of $[M]^+ = 734.9$ m/z corresponds to $[M]^+ = 721.0$ m/z, and two methyl esters corresponds to 707.1 m/z.

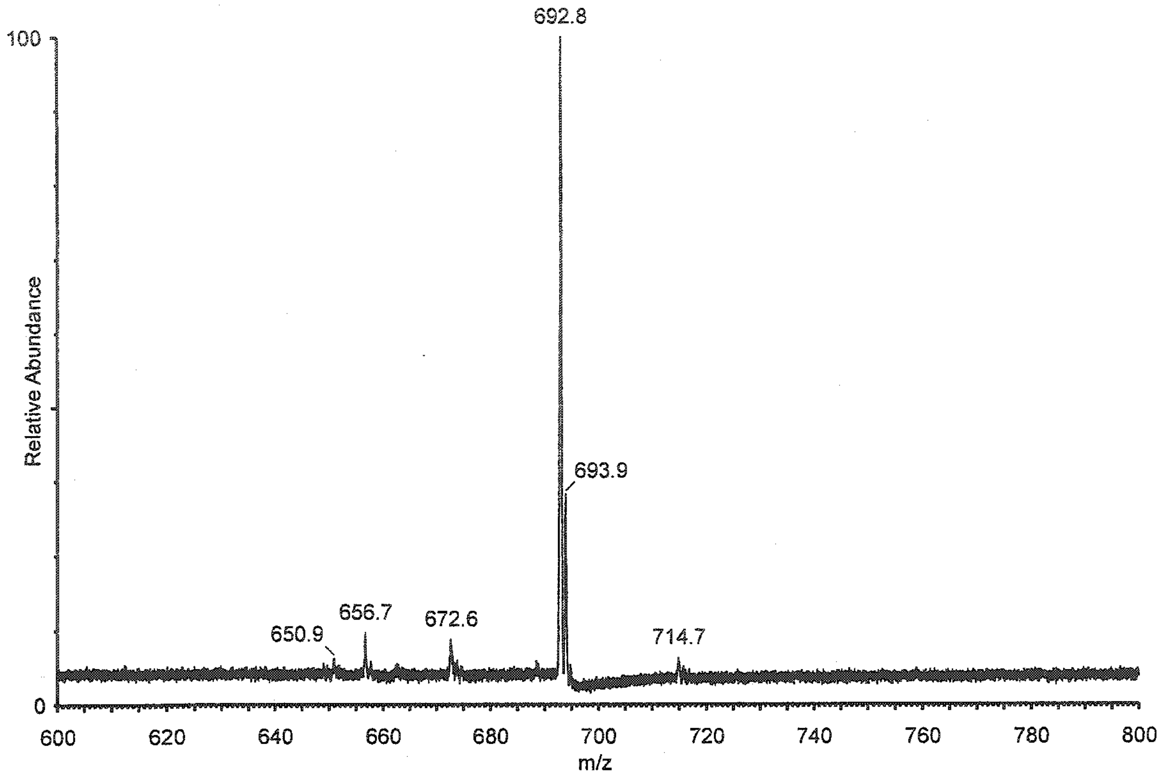


Figure 1-7 MALDI MS of an equal mixture of H-FLEEL-OH $[M + H]^+ = 650.9$, $[M + Na]^+ = 672.6$ m/z and $(CH_3)_3N^+$-FLEEL-OH $[M]^+ = 692.8$, $[M + Na]^+ = 714.7$ m/z

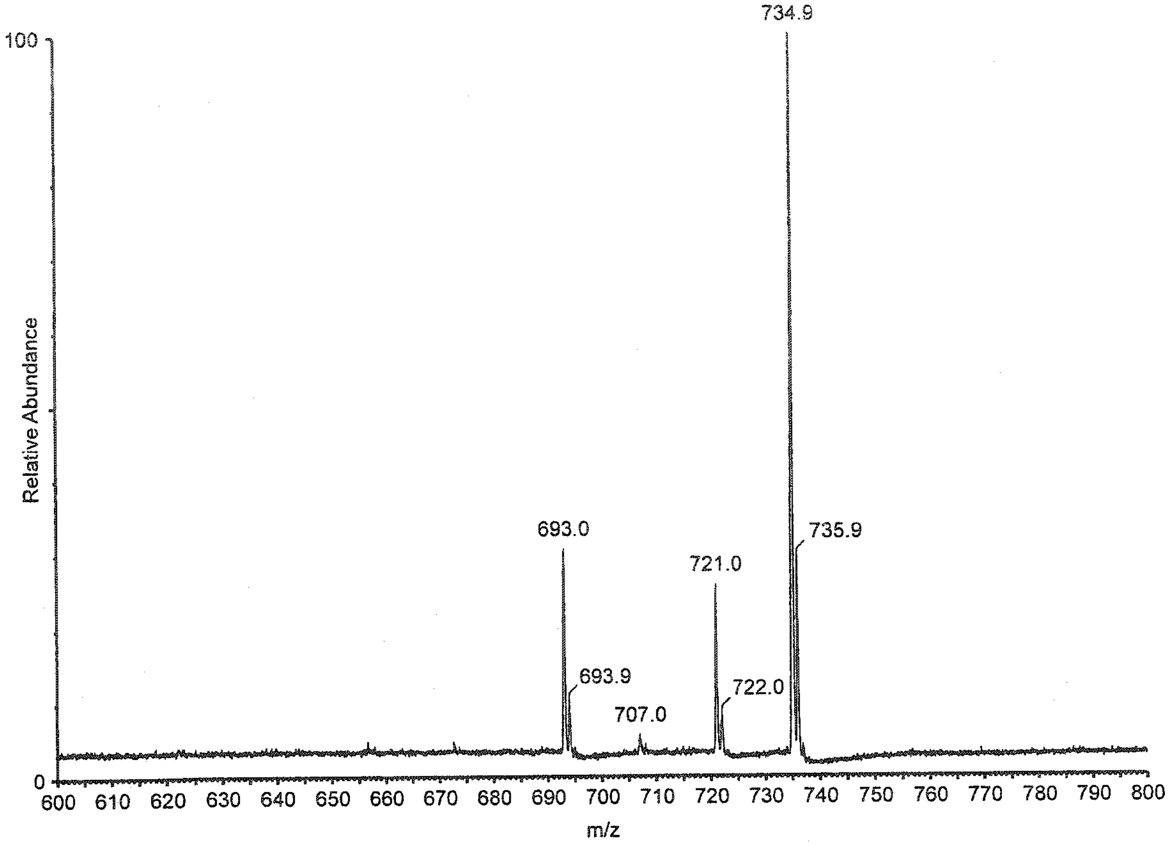


Figure 1-8 MALDI MS of an equal mixture of $(CH_3)_3N^+$-FLEEL-OH $[M]^+$= 693.0 m/z and $(CH_3)_3N^+$-FLE(OCH$_3$)E(OCH$_3$)L-OCH$_3$ $[M]^+$= 734.9 m/z

Partial hydrolysis of one methyl ester of $[M]^+$= 734.9 m/z corresponds to $[M]^+$= 721.0 m/z, and two methyl esters corresponds to 707.0 m/z.

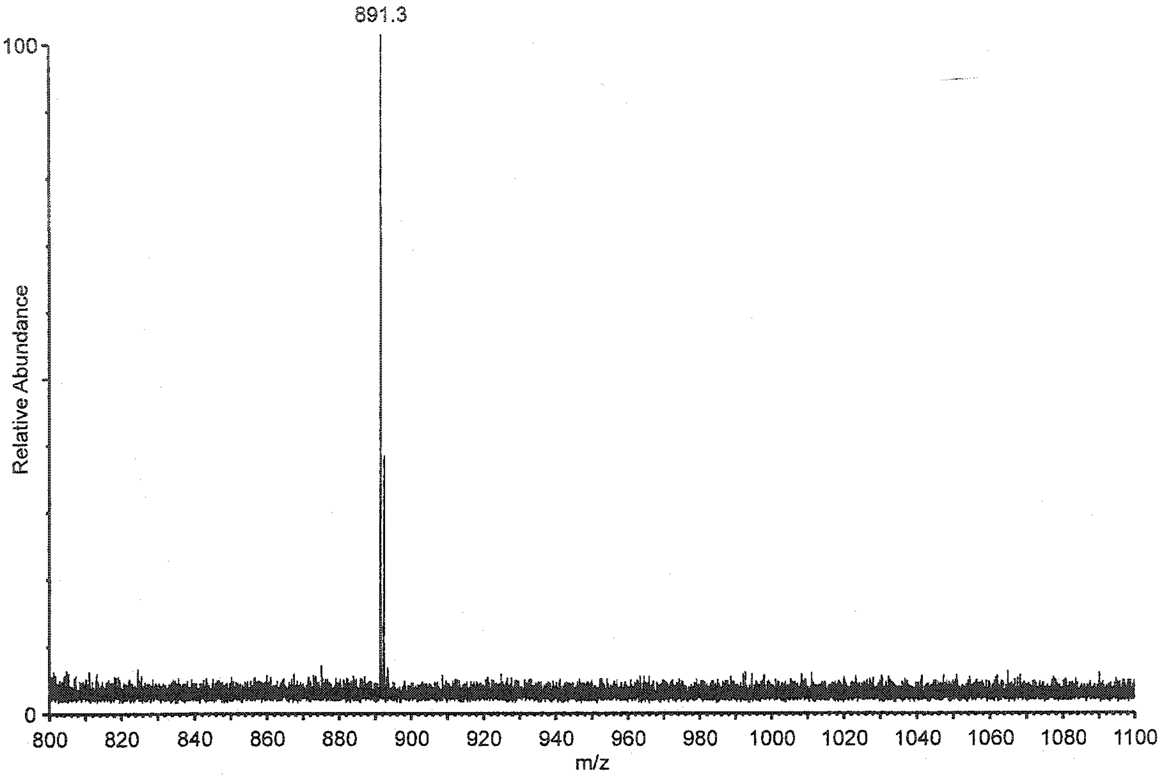


Figure 1-9 MALDI of H-GFDLNGGGVG-OH $[M + H]^+ = 891.3$ m/z

Similarly, the decapeptide H-GFDLNGGGVG-OH showed a dramatic increase in signal intensity by trimethylation of its amino group. Figure 1-9 shows the MALDI mass spectrum of this peptide alone. Unfortunately the peak observed at 891.3 m/z for the underivatized peptide H-GFDLNGGGVG-OH is 1.1 mass units lower than the expected mono-isotopic mass of 892.4 m/z for $[M + H]^+$. This discrepancy may be due to poor calibration of the mass spectrometer *e.g.* being calibrated with respect to the "mass average" value of standards and not to their mono-isotopic masses. The underivatized peptide was treated in a similar fashion as that of the methylated peptide (*i.e.* heated *in vacuo* for the same time) and a potential chemical modification which may lead to a mass difference during this step could be the deamidation of Asn-5. Deamidation of Asn-5 to Asp would yield a peptide with a gain of 1 mass unit and not the loss of 1 mu. As deamidation has not been observed in the past this was ruled out. In fact, there is no evidence from the test peptides that deamidation of Asn or Gln occurs during the *in vacuo* methylation procedure for the simple fact that deamidation would lead to the production of a free carboxyl group which, under *in vacuo* methylation, would readily produce the methyl ester. So far, no such product has been observed. Deamidation and other degradative reactions such as β-elimination at cysteines and/or serines, or peptide cleavage at susceptible sites, *i.e.* Asp-Gly, may occur during the removal of methyl esters. Hydrolysis of methyl esters was performed to simplify the spectra and to isolate the effect of trimethylation of the amino group alone on the increased detection. Hoffman elimination of the trimethylammonium group can occur under severe alkaline conditions. Although the conditions used were not severe, if this reaction occurred it would lead to the loss of $N(CH_3)_3$ from the trimethylated ammonium derivative of the peptide corresponding to a loss of 59.1 mu in the MS which was not observed. However, the permanent positive charge of the peptide would also be lost and consequently the increase in detection would be lost and if no other protonated sites exist the resulting peptide would be neutral and would not be observed in the MALDI mass spectrum. Figure 1-10 shows the MALDI mass spectrum of an equal mixture of H-GFDLNGGGVG-OH $[M + H]^+$= 891.3 m/z and $(CH_3)_3N^+$-GFD(OCH_3)LNGGGVG-OCH_3 $[M]^+$= 962.5 m/z expected (962.1 observed, discrepancy explained *vide supra*). The major peak observed at 947.1 m/z is that of the methylated peptide possessing 4 methyl groups (see Figure 1-15), which suggests that partial ester hydrolysis has occurred. Nevertheless, a dramatic increase in MALDI signal is observed for these trimethylammonium derivatives. Peaks at 916.1 and 930.1 m/z do not correspond to any methylated derivative of H-GFDLNGGGVG-OH. They may be due to impurities whose detection by MALDI have been increased by trimethylation of an amino group (*vide infra*).

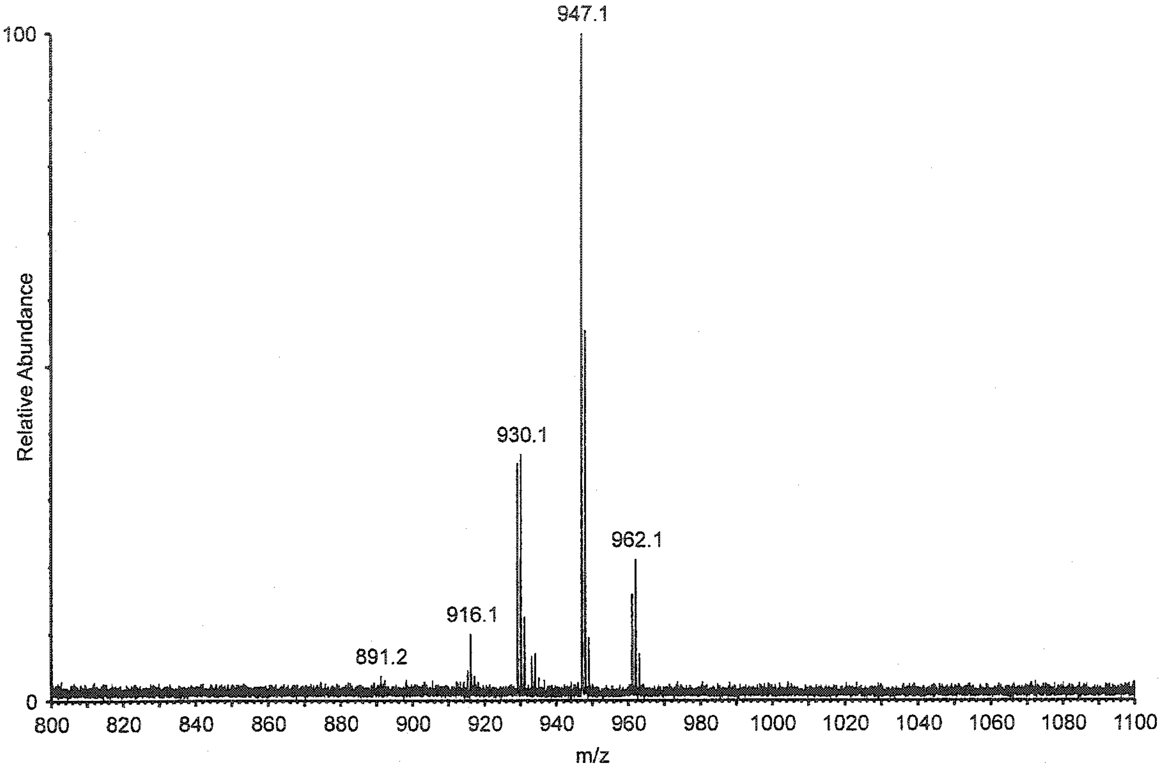


Figure 1-10 MALDI MS of an equal mixture of H-GFDLNGGGVG-OH $[M + H]^+ =$ 891.2 m/z and $(CH_3)_3N^+$-GFD(OCH_3)LNGGGVG-OCH_3 $[M]^+ =$ 962.1 m/z

Major peak at 947.1 m/z corresponds to $(CH_3)_3N^+$-GFDLNGGGVG-OCH_3 or $(CH_3)_3N^+$-GFD(OCH_3)LNGGGVG-OH. Peaks at 916.1 and 930.1 m/z are unknown.

Use of isotopes to differentiate between peptide and non-peptide signals in the mass spectrum

In vacuo methylation with CH_3I and CD_3I was also used to identify signals in the mass spectrum arising specifically from trimethylated peptides. The MALDI mass spectrum of the methylation of H-FLEEL-OH with CD_3I is shown in Figure 1-11. A predominant peak is seen at 753.0 m/z for $[M]^+$ of $(CD_3)_3N^+$-FLE(OCD_3)E(OCD_3)L-OCD_3

with some ester hydrolysis products (-17 mu). Figure 1-12 shows the MALDI mass spectrum of an equal mixture of $(CH_3)_3N^+$-FLE(OCH_3)E(OCH_3)L-OCH_3 and $(CD_3)_3N^+$-FLE(OCD_3)E(OCD_3)L-OCD_3. Some hydrolysis products are also observed however, as expected, two predominant peaks separated by 18 mu with relatively equal intensities at 734.8 and 752.8 m/z correspond to these fully methylated peptides. The removal of methyl esters greatly simplifies the spectrum, as shown in Figure 1-13. Two peaks of high intensity with a mass difference of 9 are clearly seen.

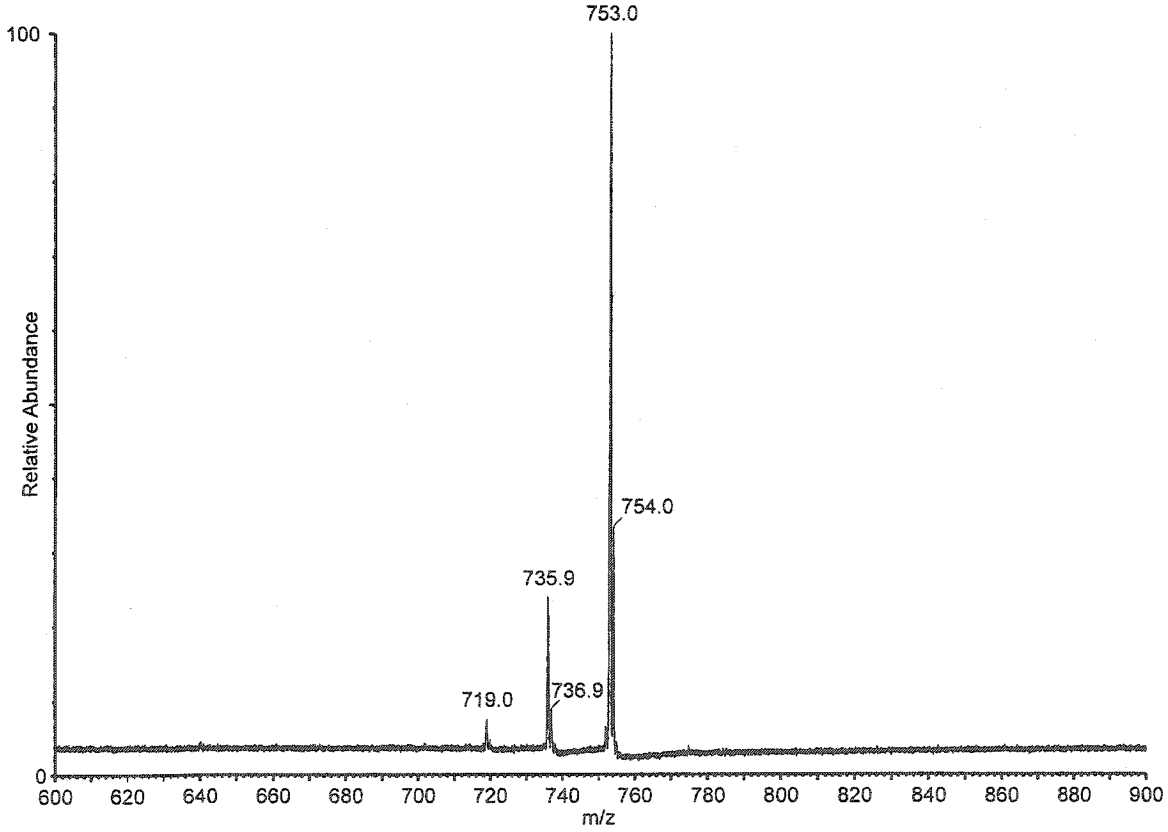


Figure 1-11 MALDI MS of $(CD_3)_3N^+$-FLE(OCD_3)E(OCD_3)L-OCD_3 $[M]^+ = 753.0$ m/z
Partial hydrolysis of one methyl ester (-17 mu) of $[M]^+ = 753.0$ m/z corresponds to $[M]^+ = 735.9$ m/z, and two methyl esters corresponds to 719.0 m/z.

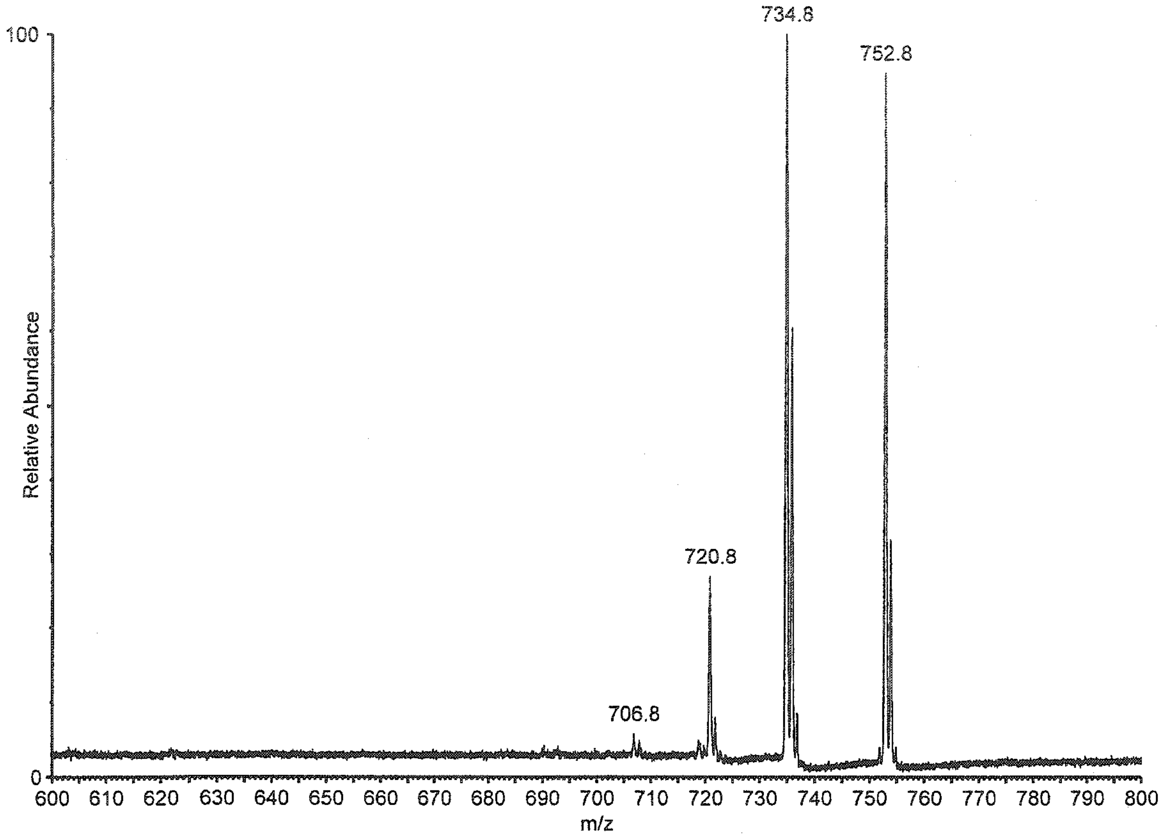


Figure 1-12 MALDI MS of an equal mixture of $(CH_3)_3N^+$-FLE(OCH_3)E(OCH_3)L-OCH_3 $[M]^+$= 734.8 m/z and $(CD_3)_3N^+$-FLE(OCD_3)E(OCD_3)L-OCD_3 $[M]^+$= 752.8 m/z

Partial hydrolysis products observed at 720.8 and 706.8 m/z.

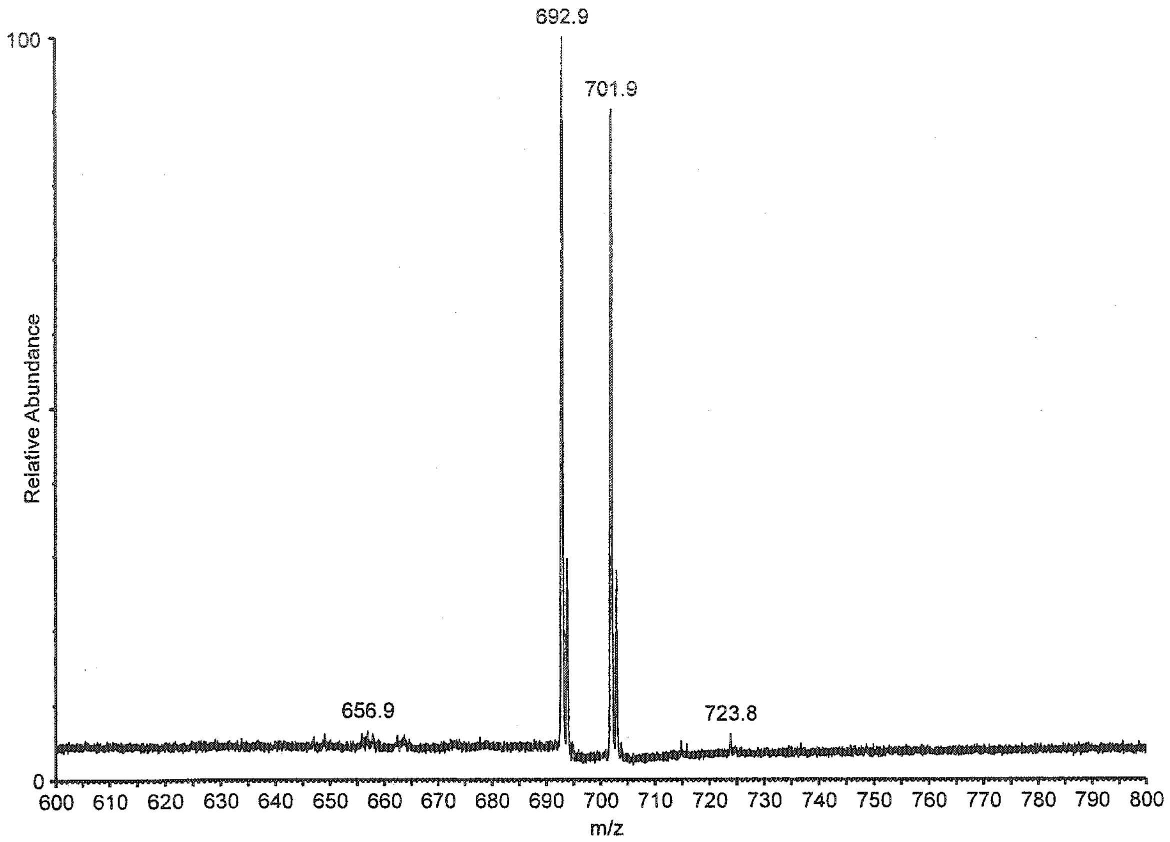


Figure 1-13 MALDI MS of an equal mixture of $(CH_3)_3N^+$-FLEEL-OH $[M]^+$= 692.9 m/z and $(CD_3)_3N^+$-FLEEL-OH $[M]^+$= 701.9 m/z, $[M + Na]^+$= 723.8 m/z

Figure 1-14 shows the MALDI mass spectrum for an equal mixture of $(CH_3)_3N^+$-GFD(OCH_3)LNGGGVG-OCH_3 $[M]^+$= 961.1 m/z (observed) and $(CD_3)_3N^+$-GFD(OCD_3)LNGGGVG-OCD_3 $[M]^+$= 976.2 m/z. As expected, two predominant peaks of near equal intensity, with a 15 mu difference, are observed with the presence of some methylated impurities. Figure 1-15 shows the MALDI mass spectrum of what was expected to be an equal mixture of $(CH_3)_3N^+$-GFDLNGGGVG-OH ($[M]^+$= 934.4 m/z) and $(CD_3)_3N^+$-GFDLNGGGVG-OH ($[M]^+$= 943.4 m/z) (after ester hydrolysis) to give two peaks separated by 9 mu. While these peaks are present, two major peaks instead are observed at 947.2 and 959.3 m/z corresponding to a mass difference of 12.1 which suggests the presence of 4 methyl groups on each peptide therefore one of two methyl esters remaining. There is also a peak at 930.2 m/z which could be the result of the loss of a deuterated methyl group (-17 mu) however this cannot be the case as this peak is also observed in Figure 1-10 for the non-deuterated sample. This peak could arise from the trimethylammonium derivative of an unobservable peptide impurity. The mass also corresponds to the potassium adduct of the underivatized peptide however this seems unlikely as this adduct is not observed in the MALDI mass spectrum for the underivatized peptide sample.

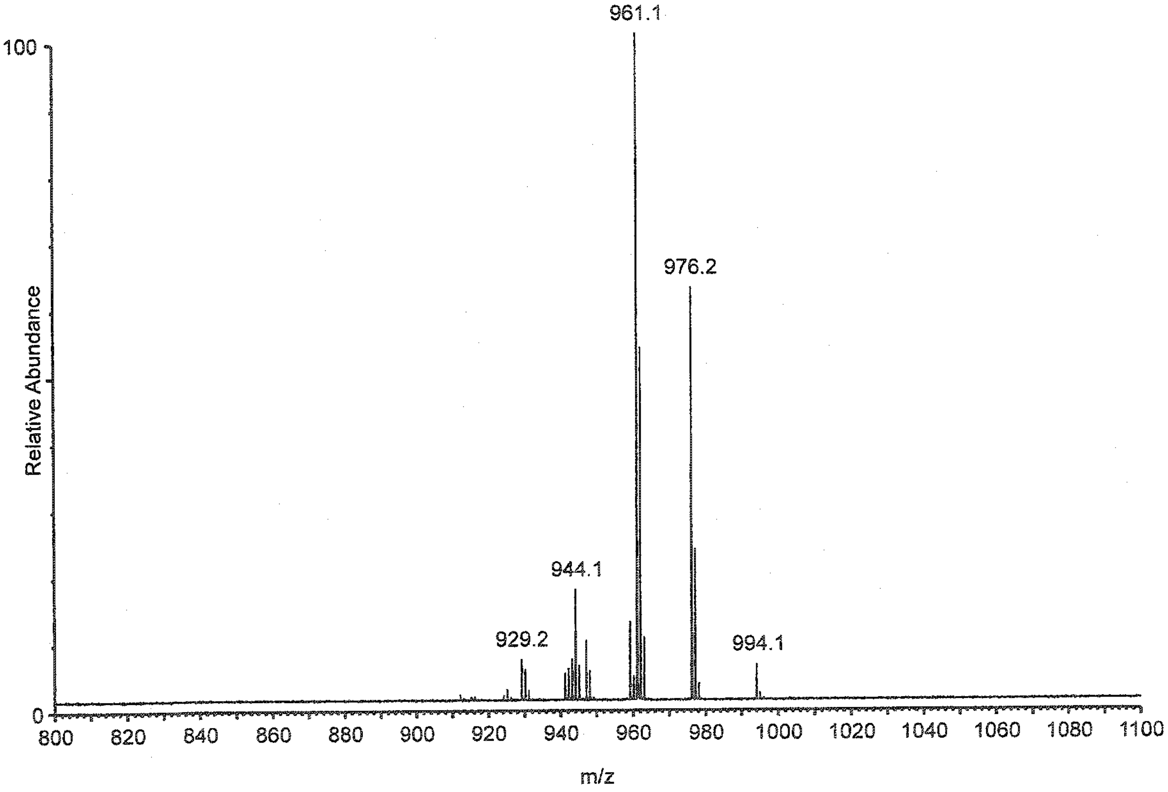


Figure 1-14 MALDI MS of an equal mixture of $(CH_3)_3N^+$-GFD(OCH_3)LNGGGVG-OCH_3 $[M]^+$= 961.1 m/z and $(CD_3)_3N^+$-GFD(OCD_3)LNGGGVG-OCD_3 $[M]^+$= 976.2 m/z

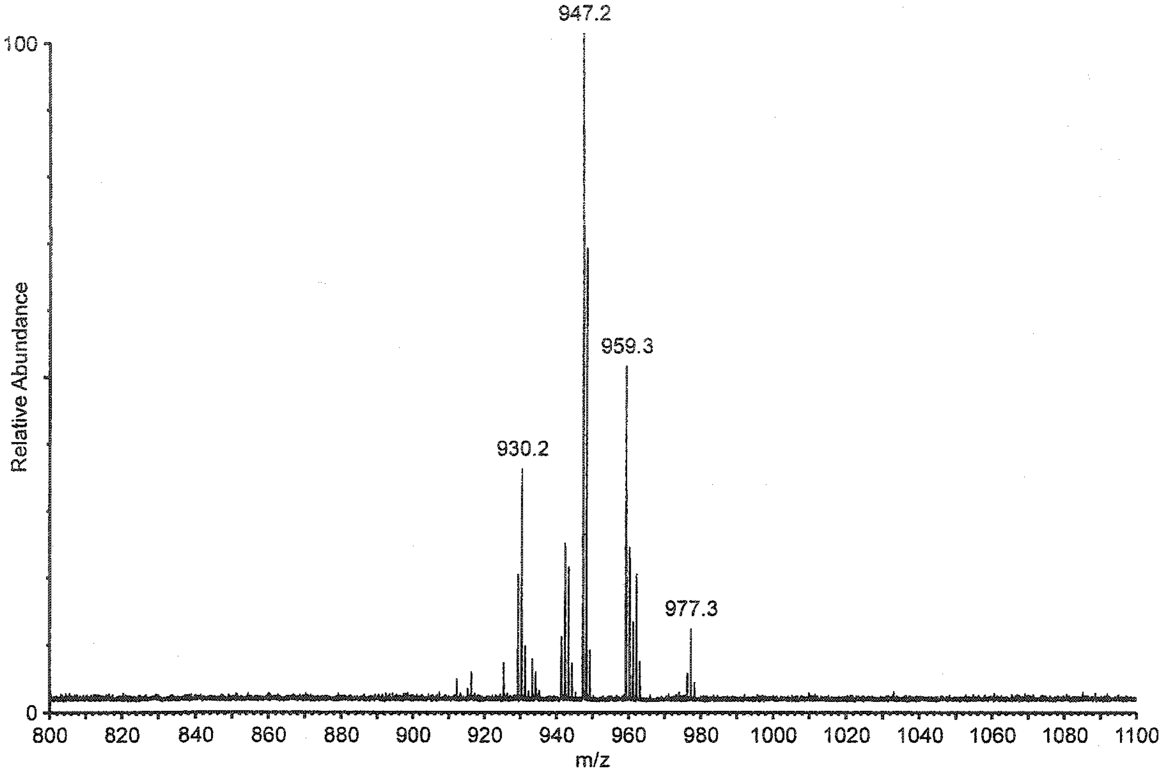


Figure 1-15 MALDI MS of an equal mixture of $(CH_3)_3N^+$-GFDLNGGGVG-OH and $(CD_3)_3N^+$-GFDLNGGGVG-OH after base hydrolysis.

Peaks at 947.2 and 959.3 m/z correspond to incomplete hydrolysed products with one methyl ester remaining. Peak at 977.3 m/z corresponds to $(CD_3)_3N^+$-GFD(OCD_3)LNGGGVG-OCD_3. Peak at 930.2 m/z from unknown impurity.

The hydrolysis of the methyl esters was not complete under these conditions, complicating the mass spectrum and, in fact, in one instance after the removal of the methyl esters by base hydrolysis a significant peak was observed at approximately double the mass (data not shown). This may have originated from the coupling between peptides via the methyl ester at the unhindered N and C-terminal glycines. A means to simplify this would be to completely esterify the peptide either *in vacuo* (Vakos *et al.*, 2001) or in solution, *e.g.* with MeOH-HCℓ (Goodlett *et al.*, 2001) instead of hydrolyzing them. The use of deuterated and non-deuterated methanol in the presence of an acid catalyst (*e.g.* HCℓ, BF_3 or acetyl chloride to produce HCℓ *in situ*) could be used in both methods as the reagents are volatile. This method would be similar to the methylation procedure for the analysis of ritalinic acid in urine with BF_3-MeOH for gas chromatographic analysis (Vu-Duc and Vernay, 1992). Non-isotopically labelled methanol could be employed to greatly simplify the spectrum if isotopically labelled iodomethane was initially used to label the peptides as all methyl esters produced during iodomethane methylation would completely exchange. The resulting derivatized peptide would display a doublet in the mass spectrum with a mass difference solely dependent on the number methyl groups incorporated from iodomethane.

The mass spectrum of H-VGVAPG-OH has a signal with the expected mass at 499.0 m/z but several other peaks, including one of higher mass at 567.6 m/z with a greater intensity than the peptide, are also present, see Figure 1-16.

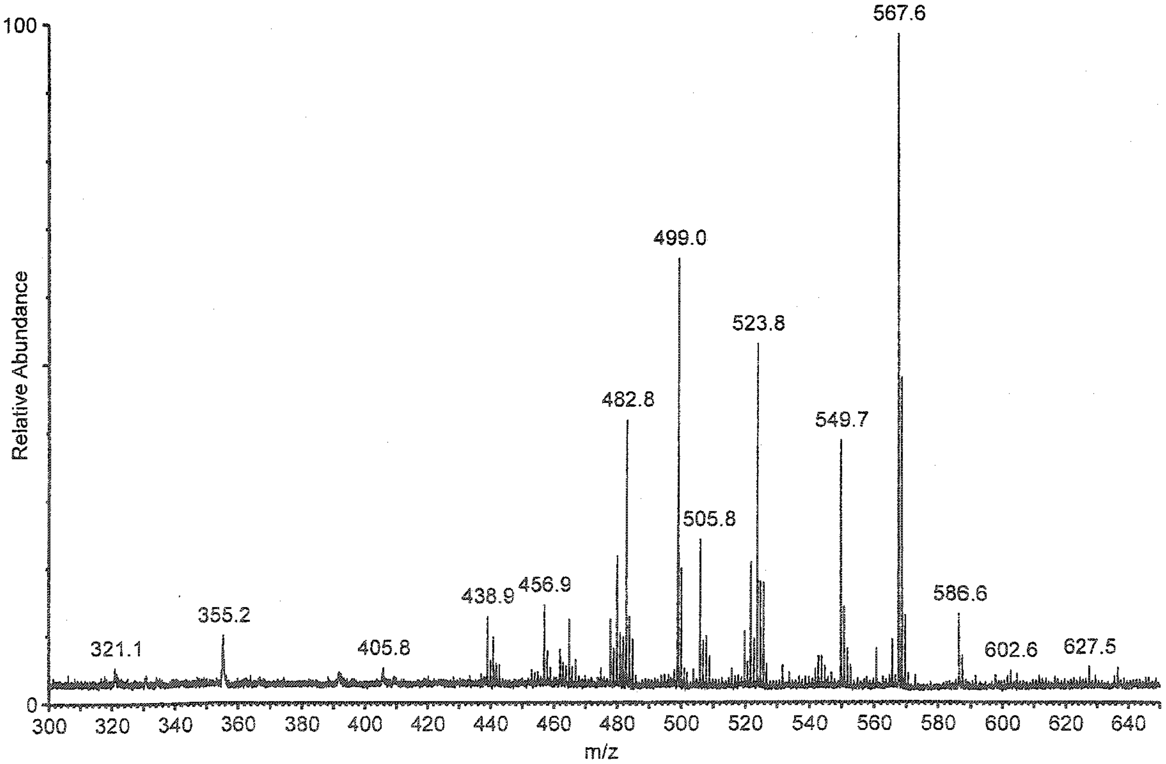


Figure 1-16 MALDI MS of H-VGVAPG-OH $[M + H]^+ = 499.0$ m/z

Without additional information, it is not possible to identify which of these signals corresponds to a peptide let alone to the peptide of interest. The mass spectrum of a mixture of equal amounts of H-VGVAPG-OH and $(CH_3)_3N^+$-VGVAPG-OH, gives the expected signal forty-two mass units higher at 541.0 m/z. Again, a large increase in signal intensity is observed for the trimethyl α-amino peptide Figure 1-17.

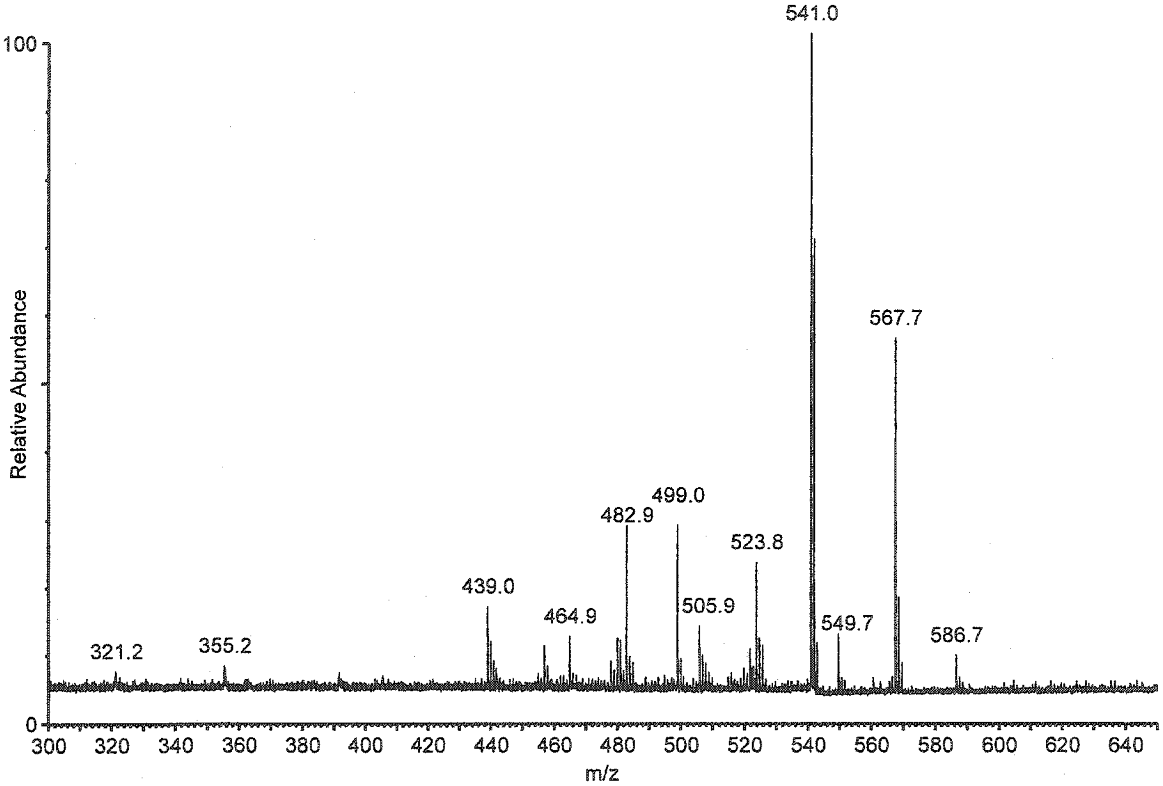


Figure 1-17 MALDI MS of an equal mixture of H-VGVAPG-OH $[M + H]^+$= 499.0 m/z and $(CH_3)_3N^+$-VGVAPG-OH $[M]^+$= 541.0 m/z

When equal amounts of base hydrolysed CH_3I and CD_3I methylated H-VGVAPG-OH are combined, two predominant peaks are observed at 541.0 and 550.0 m/z (see Figure 1-18), confirming that the signal at 499.0 m/z in Figure 1-16 is from H-VGVAPG-OH and that the other signals, including the strongest one at 567.7 m/z, are due to non-peptide impurities which do not possess an amino group.

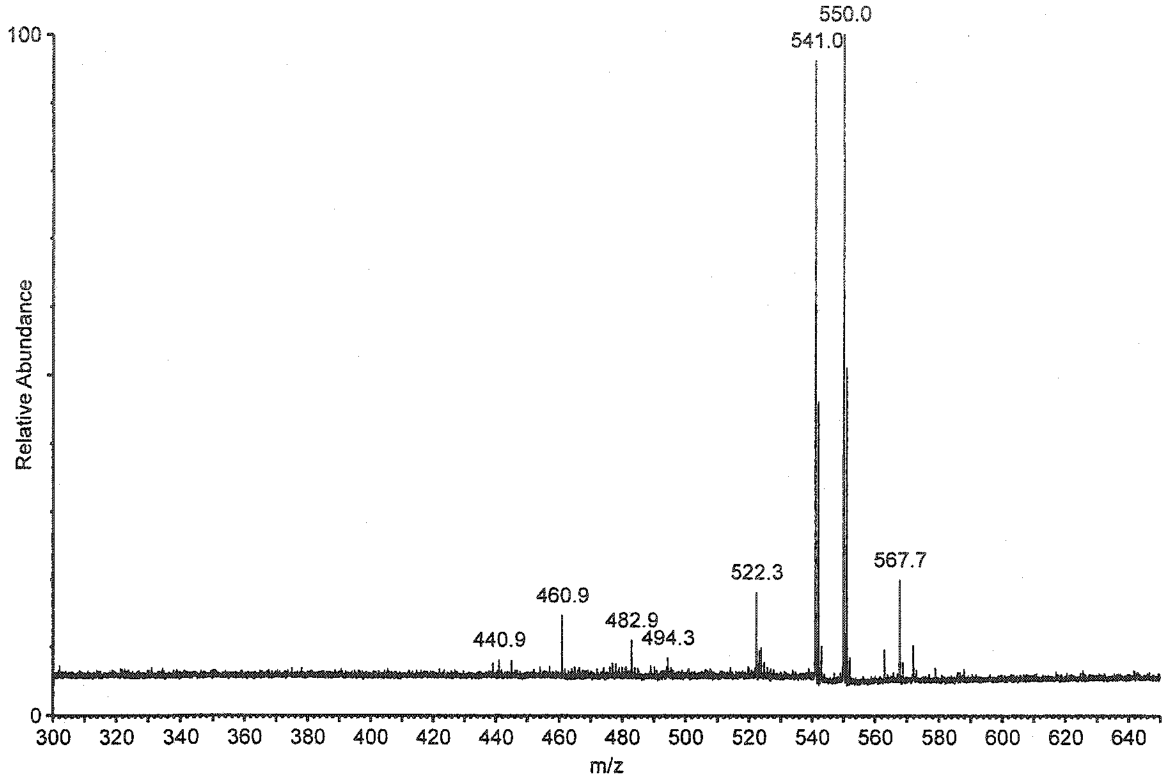


Figure 1-18 MALDI MS of an equal mixture of $(CH_3)_3N^+$-VGVAPG-OH $[M]^+= 541.0$ m/z and $(CD_3)_3N^+$-VGVAPG-OH $[M]^+= 550.0$ m/z

For the peptide: H-GFDLNGGGVG-OH, in order to compare the ionization yield of the trimethylammonium derivative to the underivatized peptide, a relatively large amount (5 pmol) was applied to the sample plate for MALDI MS. The methylated derivative ($[M]^+= 962.3$) is easily seen at the 5 pmol level compared to the unmodified peptide ($[M + H]^+= 892.1$) (Figure 1-10). Some minor impurities (see Figure 1-9), not detected in the untreated sample, are also detected after reaction with iodomethane. Since the peptide was employed as received from the supplier without further purification, it is likely that these are peptide or other amine containing impurities whose detection is also enhanced by reaction with iodomethane.

Effect of trimethylation of peptide ε-amino groups on MALDI MS detection

In order to determine the effect of trimethylating the ε-amino of lysine on the increase in detection by MALDI, lyophilized insulin B_{ox}-chain was reacted *in vacuo* with iodomethane and digested with TPCK trypsin.

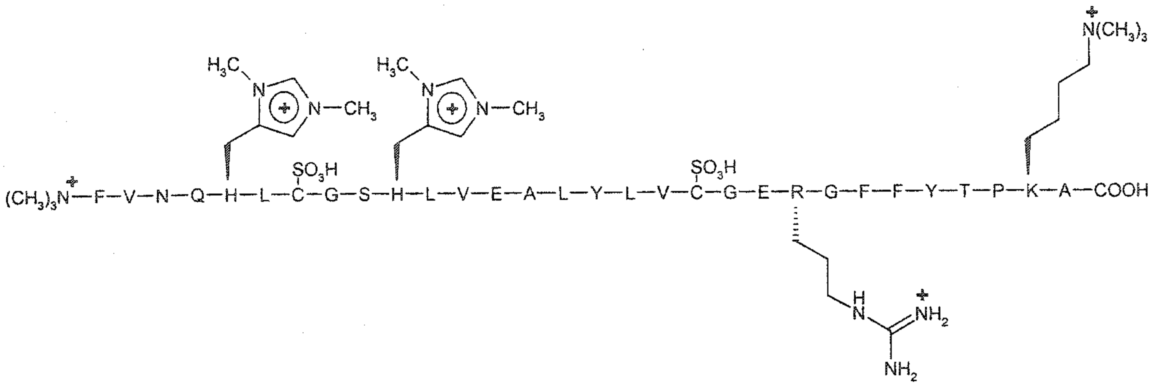


Figure 1-19 Schematic representation of the amino acid sequence of *in vacuo* methylated bovine insulin B_{ox}-chain (esters removed)

The insulin B_{ox}-chain consists of thirty amino acids with one arginine at position 22 and one lysine at position 29. The methylated peptide, after removal of methyl esters, is shown in Figure 1-19. It is predicted that trypsin digestion will generate peptide H-GFFYTPK-OH with $[M + H]^+ = 859.2$ m/z and the derivatized B_{ox}-chain will generate peptide H-GFFYTPK(ε-$N^+(CH_3)_3$)A-OH (see Scheme 1-2) with $[M]^+ = 972.2$ m/z as trypsin does not cleave peptide bonds at a trimethylated lysine (Seely and Benoiton, 1970). The results shown in Figure 1-20 and Figure 1-21 confirm that these are indeed the peptides generated by trypsin digestion.

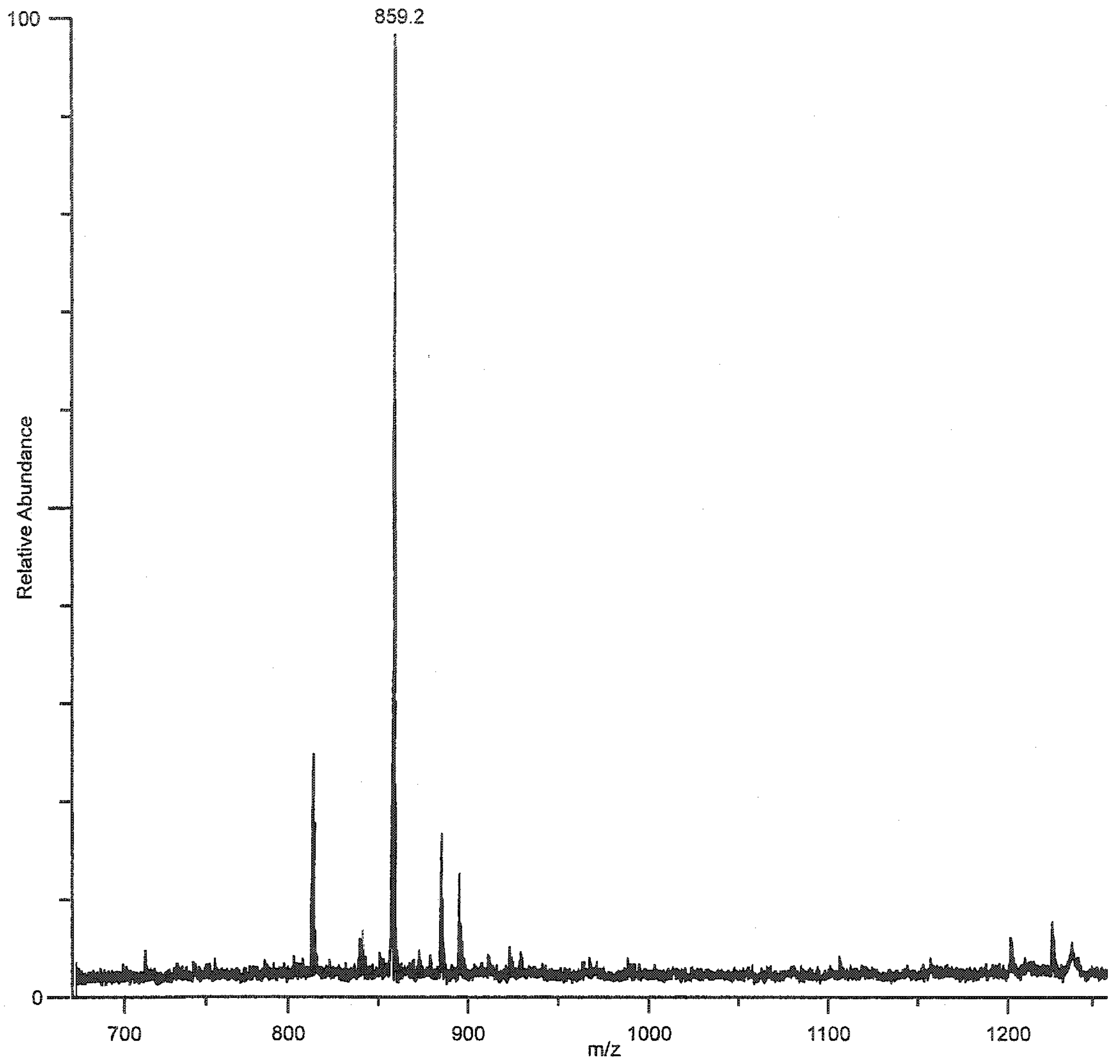


Figure 1-20 MALDI MS of tryptic peptides of insulin B_{ox}-chain

When equal amounts of H-GFFYTPK-OH, $[M + H]^+ = 859.2$ m/z, and H-GFFYTPK($\varepsilon N^+(CH_3)_3$)A-OH, $[M + H]^+ = 972.4$ m/z, are subjected to MALDI MS, the peptide with the trimethylated ε-amino group gives a much greater signal intensity (Figure 1-21).

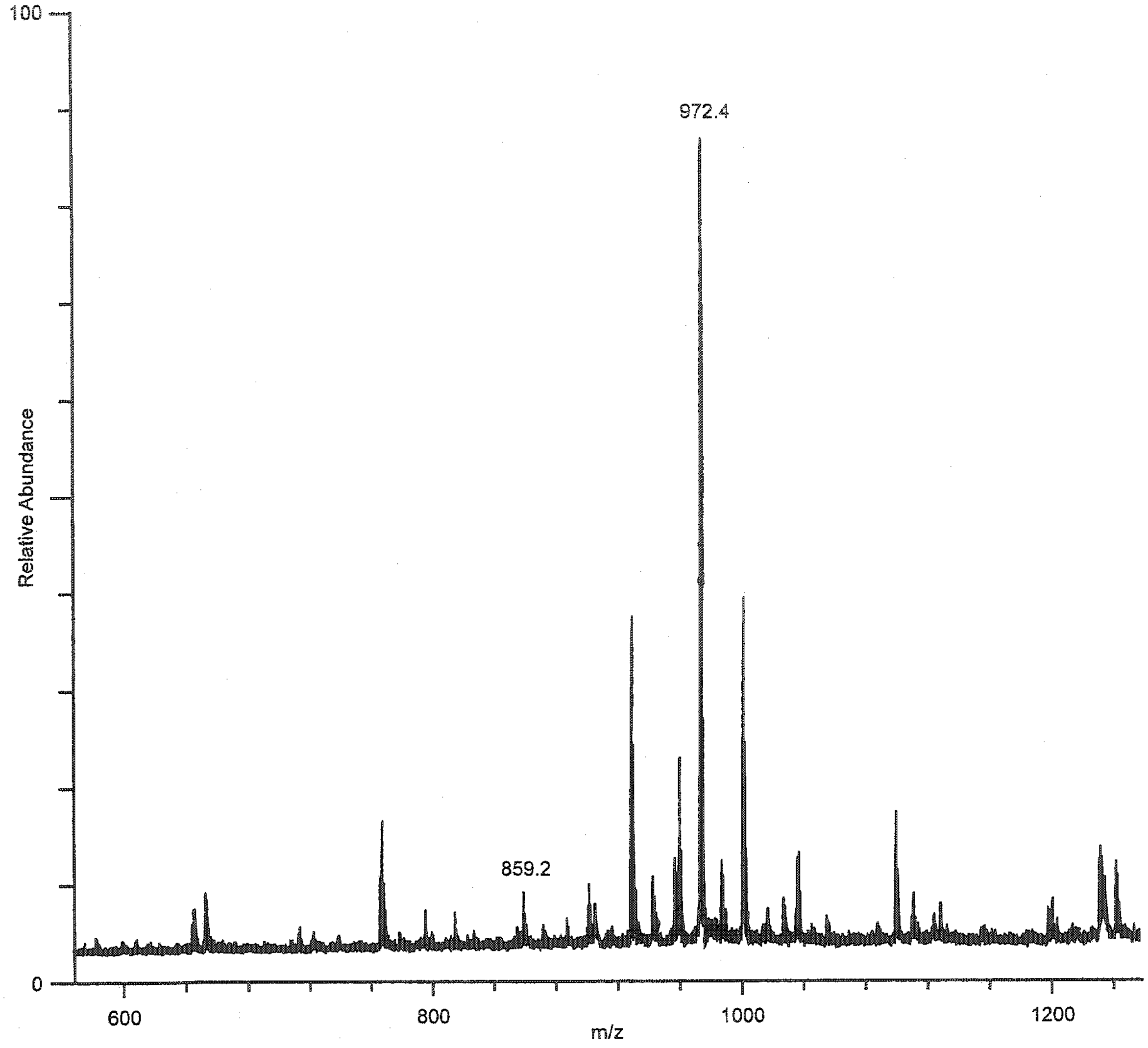


Figure 1-21 MALDI MS of the tryptic peptides generated from an equal mixture of insulin B_{ox}-chain and methylated B_{ox}-chain

If equal amounts of insulin B_{ox}-chain are reacted with ^{12}C and ^{13}C iodomethane, digested with trypsin and mixed together, it is expected that the trimethyl-ϵ-amino lysine peptide will show two signals three mass units apart. This is shown in the expanded spectrum in Figure 1-22 where the expected signals at 972.4 and 975.4 m/z are observed confirming that a trimethylated peptide is present. Similarly the signals at 928.0 and 931.0 m/z indicate the presence of another trimethylated peptide. However, the mass does not correspond to any possible trimethylated peptide from the insulin B_{ox}-chain or any possible trimethylated peptide derived from oxidized insulin A-chain which may be present as an impurity. The other signals present are not separated by three mass units and are therefore not peptides containing a trimethylated amino group.

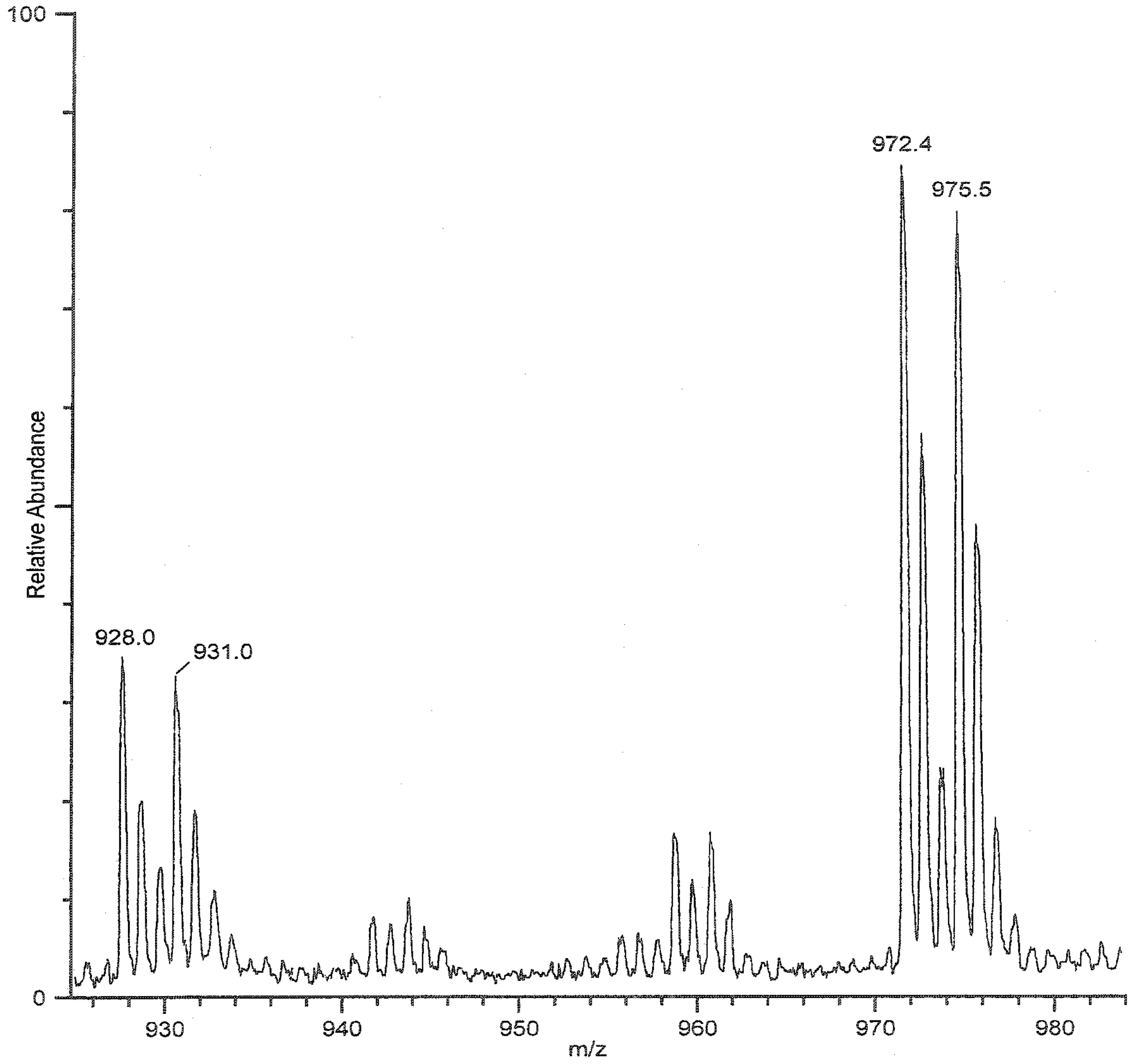
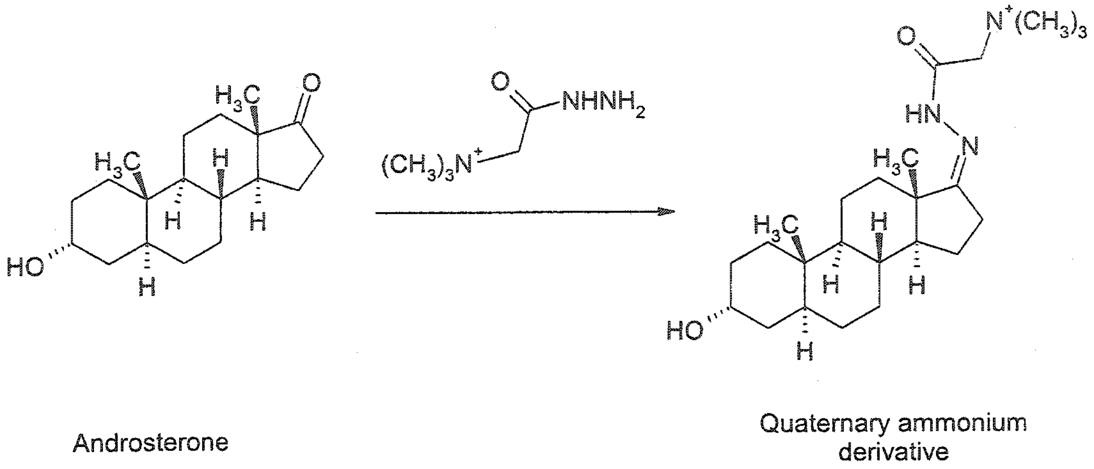


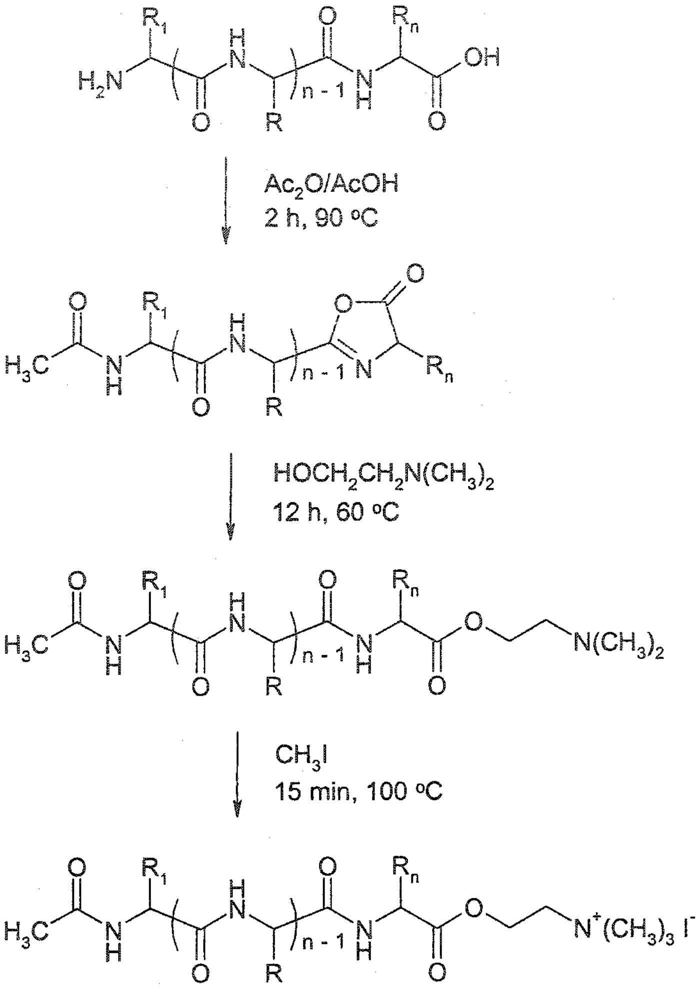
Figure 1-22 Expanded portion of the MALDI MS of the tryptic peptides generated from an equal mixture of insulin B_{ox}-chain methylated with ^{12}C and ^{13}C iodomethane

An example of the introduction of a charged group into a non peptide molecule to enhance its detection has been found in the use of fast atom bombardment mass spectrometry (FAB MS). The ketosteroid androsterone is undetectable under normal FAB conditions but gives an intense molecular ion peak in the FAB MS spectrum after derivatization to a quaternary ammonium derivative with Girard's reagent T (see Scheme 1-3) (Knapp, 1990).



Scheme 1-3 Derivatization of androsterone to a quaternary ammonium derivative with Girard's reagent T (from Knapp, 1990)

Conversely, Biemann and co-workers, reported that introduction of a trimethylated quaternary ammonium group at the C-terminus of a peptide, *viz.* *N*-acetyl-Leu-enkephalin trimethylammonium ethyl ester (see Scheme 1-4), did **not** appear to improve its signal intensity with FAB MS (Vath *et al.*, 1988). Their result showing no improvement in the signal intensity may have been due to incomplete derivatization as the method involves three steps to produce the trimethyl peptide derivative.



Scheme 1-4 Conversion of Leu-enkephalin to a C-terminal quaternary ammonium derivative (from Vath *et al.*, 1988)

Larger amounts than necessary of trimethylated peptide were used to obtain the MALDI MS spectra, so that comparisons could be made with the unmodified peptides. The results clearly show that trimethylated peptides are readily detectable at this level. Based on the "pH memory" effect observed for lyophilized proteins, in which the pH of the solution from which the protein was prepared is "remembered" with no proton exchange (Zaks and Klibanov, 1988; Xu and Klibanov, 1996; Klibanov, 1997), one can assume that not all of the α or ε amino groups of the peptides are trimethylated under the conditions employed (Kaplan and Taralp, 1997; Taralp and Kaplan, 1997; Vakos *et al.*, 2000). Consequently the increase in signal intensity observed is likely greater than that indicated. Lyophilization at lower pH values was used to minimize the possibility of peptide modification or salt effects. To obtain complete trimethylation of amino groups, reaction with iodomethane can be carried out on peptides lyophilized at higher pH values. After *in vacuo* methylation, the trimethylated peptides were loaded onto the target without any purification such as micro reverse phase columns that are commonly employed to remove salts. While a small amount of salt is present due to the *in vacuo* methylation procedure, such treatment does not appear to be necessary, but in some cases it may improve the quality of the spectra obtained.

In order to simplify interpretation of the data the peptides tested did not possess free sulfhydryls or disulfide bonds. Peptides with free sulfhydryls or disulfide bonds would most likely be cleaved under the alkaline aqueous conditions by the heat induced β-elimination of cysteine residues (Kim and Kim, 2001). This cleavage would result in a triplet in the mass spectrum corresponding to the dehydroalanine, cysteine and thiocysteine containing peptide derivatives.

Application of the *in vacuo* methylation for the identification of peptides from enzymatic digests of proteins

The use of isotope mixtures clearly helps in discriminating between mass peaks derived from peptide and those arising from impurities in the MALDI mass spectrum as seen in Figure 1-22. Its use to identify peptides from enzymatic digests of *in vacuo* methylated proteins was investigated. Figure 1-23 shows the MALDI mass spectrum of a papain digest of an equal mixture of BSA and CH_3 methylated BSA. Numerous signals are observed, many of which have strong intensities.

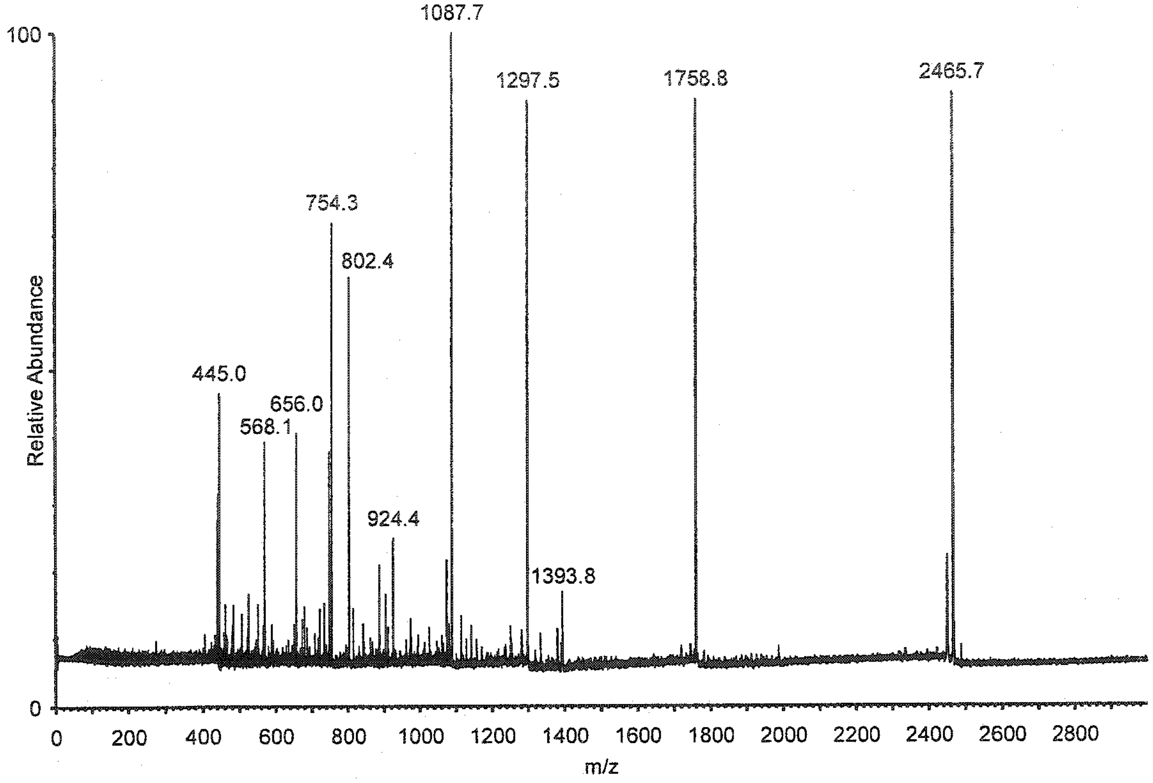


Figure 1-23 MALDI MS of the papain digest of an equal mixture of BSA and CH_3 *in vacuo* methylated BSA

From an equal mixture of papain digested, CH_3 and CD_3-methylated BSA, only one very strong doublet (highest signal in the mass spectrum) of relatively equal intensity 6 mu apart at 1087.8 and 1093.8 m/z is apparent (see Figure 1-24 and expanded portion in Figure 1-25). This does not match any possible N-terminal derived peptide, instead it does match one possible peptide (200 ppm search tolerance): the peptide I[145-152]F (IARRHPYF) $[M]^+$= 1087.6 ± 0.2 m/z. This peptide contains a histidine residue which would be methylated to the dimethylimidazolium derivative which gives rise to a doublet separated by 6 mu. The 2 arginines in this peptide seem to give this peptide a high ionization yield under MALDI. It is surprising that in the mass spectrum (Figure 1-24) no other doublet is apparent, especially for peaks displaying high intensities. BSA has 25 arginines and it is suspected that the strong peaks observed in these spectra arise from arginine containing peptides. The results clearly demonstrate the difficulty encountered with differential ionization yields during MALDI.

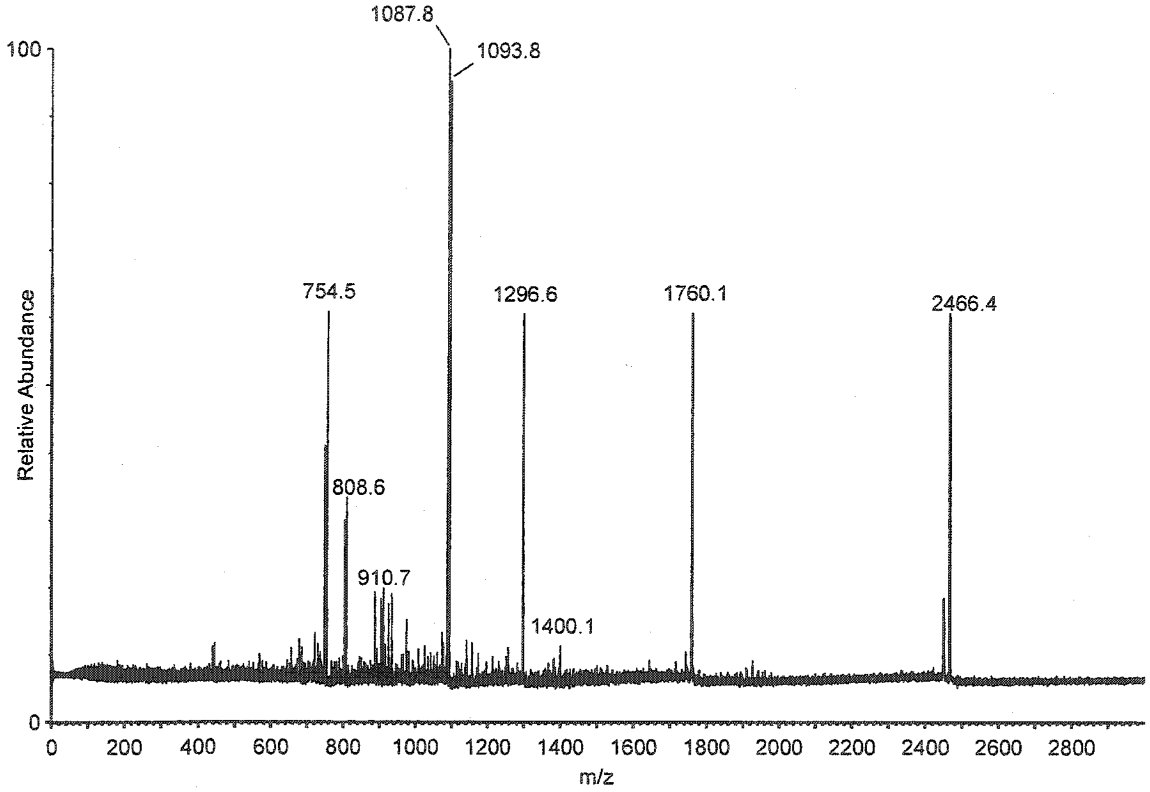


Figure 1-24 MALDI MS of the papain digest of an equal mixture of CH_3 and CD_3 *in vacuo* methylated BSA

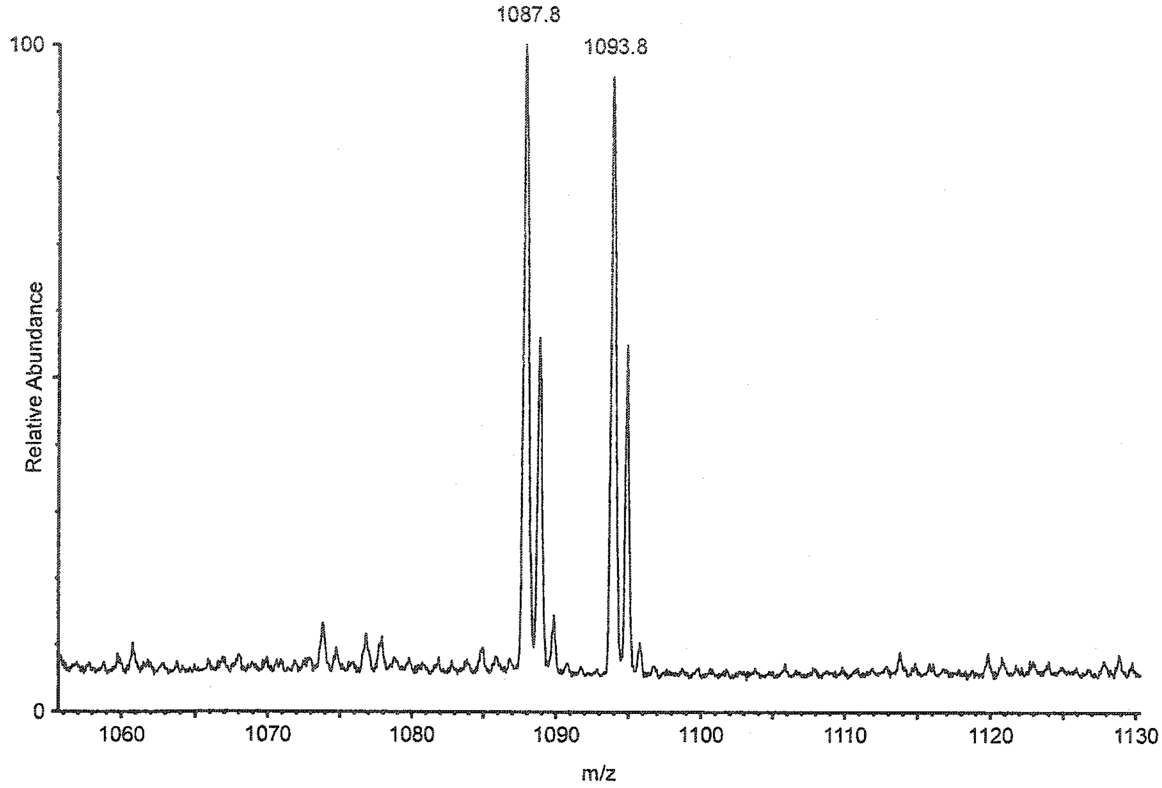


Figure 1-25 Portion of the MALDI MS of an equal mixture of peptides generated from a papain digest of CH_3 and CD_3 *in vacuo* methylated BSA

In Figure 1-26 the total papain digest of an equal mixture of CH_3 and CD_3 *in vacuo* methylated hHb (LpH 8.0) is shown. LpH 8.0 was chosen to target α-amino groups over ε-amino groups although methylation of lysines did occur as multiple peaks differing by multiples of three are visible. hHb consists of two polypeptides the α-chain and the β-chain, and *in vivo* exists as a $\alpha_2\beta_2$ globular tetramer.

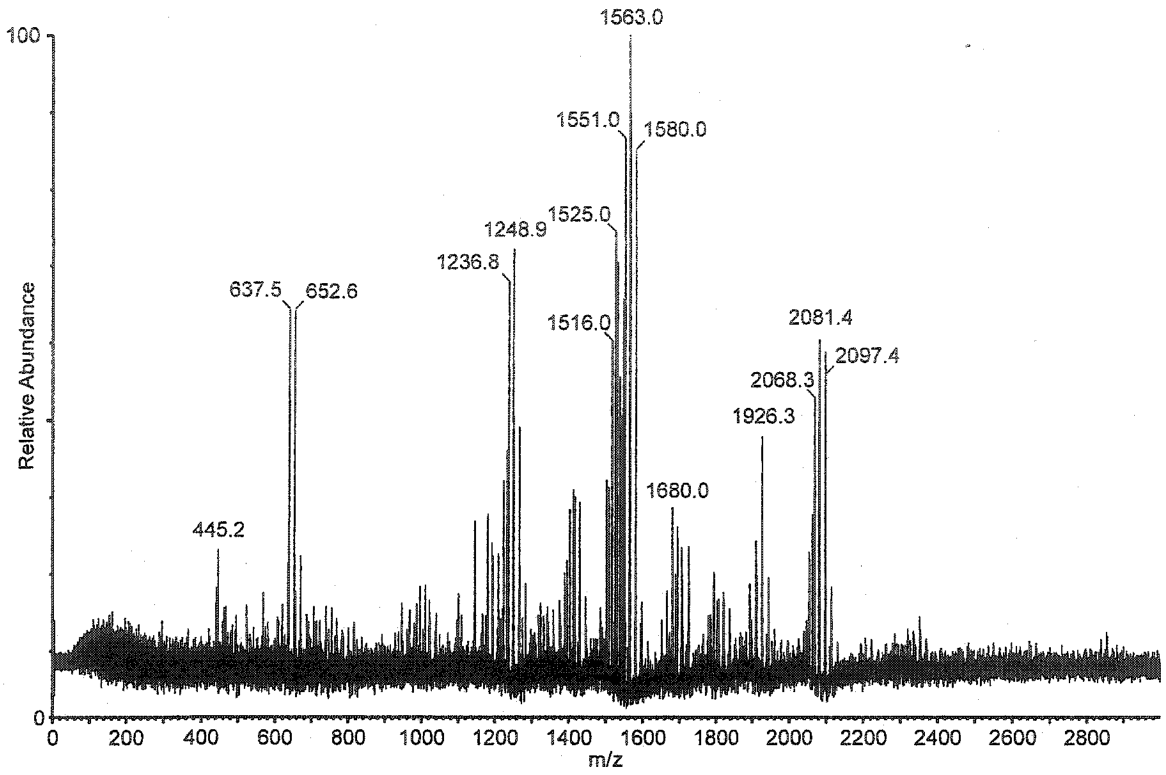


Figure 1-26 Full MALDI MS of the papain digest of an equal mixture of CH_3 and CD_3 *in vacuo* methylated hHb

Figure 1-27 shows a portion of the MALDI MS of the total papain digest of an equal mixture of CH_3 and CD_3 methylated hHb. Two intense peaks of similar intensity are observed at 1236.8 and 1248.9 m/z. These peaks are 12.1 mu apart which suggest that they originate from trimethylammonium derivatized peptides possessing a total of 4 methyl groups. The fourth methyl group is suspected to be a methyl ester as a pair of peaks, 14 and 17 mu lower (corresponding to the loss of CH_3 and CD_3, respectively), are also present in spectrum.

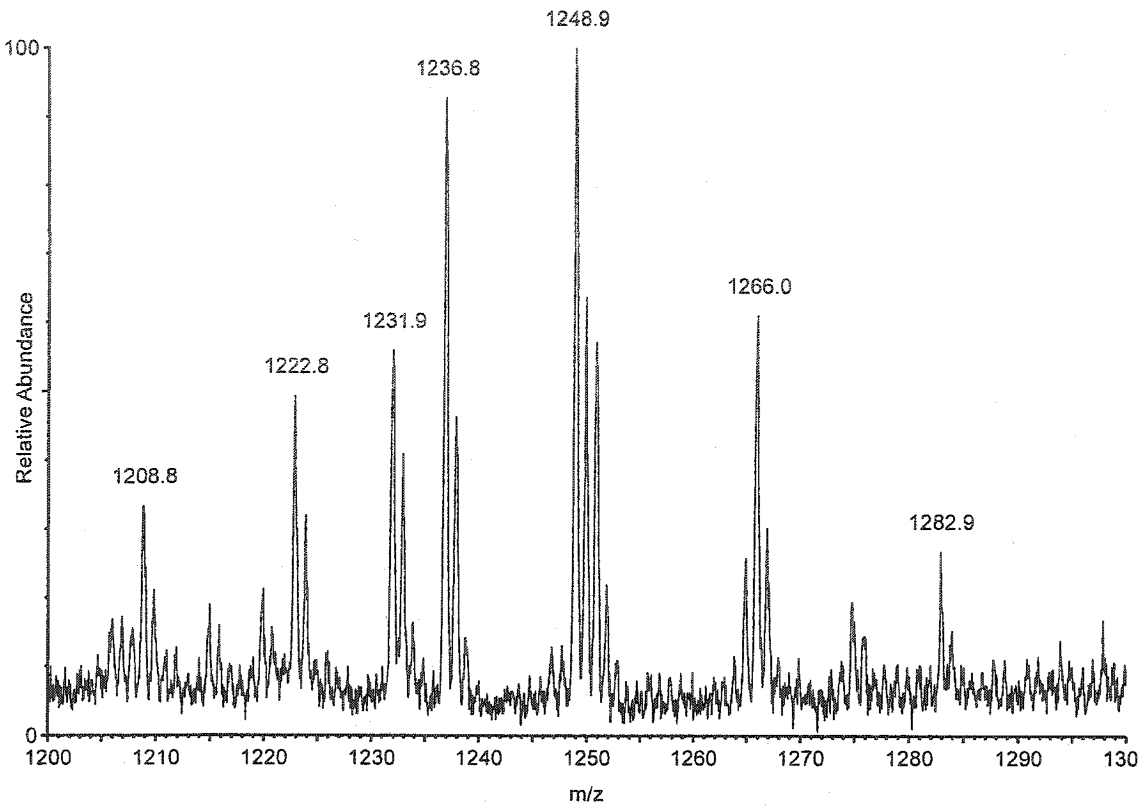


Figure 1-27 Portion of the MALDI MS of the papain digest of an equal mixture of CH_3 and CD_3 *in vacuo* methylated hHb

With a search tolerance of 100 ppm three peptides were found to fit: the peptides L[125-135]V $[M]^+ = 1236.7 \pm 0.1$ m/z (LDKFLASVSTV) or D[126-136]L $[M]^+ = 1236.7 \pm 0.1$ m/z (DKFLASVSTVL) derived from the α-chain or peptide L[81-91]L (LKGTFATLSEL) $[M]^+ = 1236.7 \pm 0.1$ m/z derived from the β-chain. Each contain a lysine and one acidic amino acid, which would give rise to two peaks 12 mass units apart corresponding to the deuterated and non-deuterated peptide pairs. The loss of the methyl ester is also observed with the corresponding peptide pair at 1222.8 and 1231.9 m/z. These results again demonstrate the increase in signal intensity by the trimethylammonium derivatization of a lysine containing peptide.

Figure 1-28 shows a portion of the MALDI mass spectrum. The peaks at 637.5 and 652.6 m/z are relatively intense and are 15.1 mu apart suggesting that these are trimethylammonium derivatized peptides containing 5 methyl groups. This is consistent with the methylated pentapeptide derived from the N-terminal of the β-chain: $(CH_3)_3N^+$-VH(*imN*$_1$(CH_3),*imN*$_3^+$(CH_3))LTP-OH possessing the trimethylammonium N-terminal derivative and the dimethylimidazolium derivative of histidine.

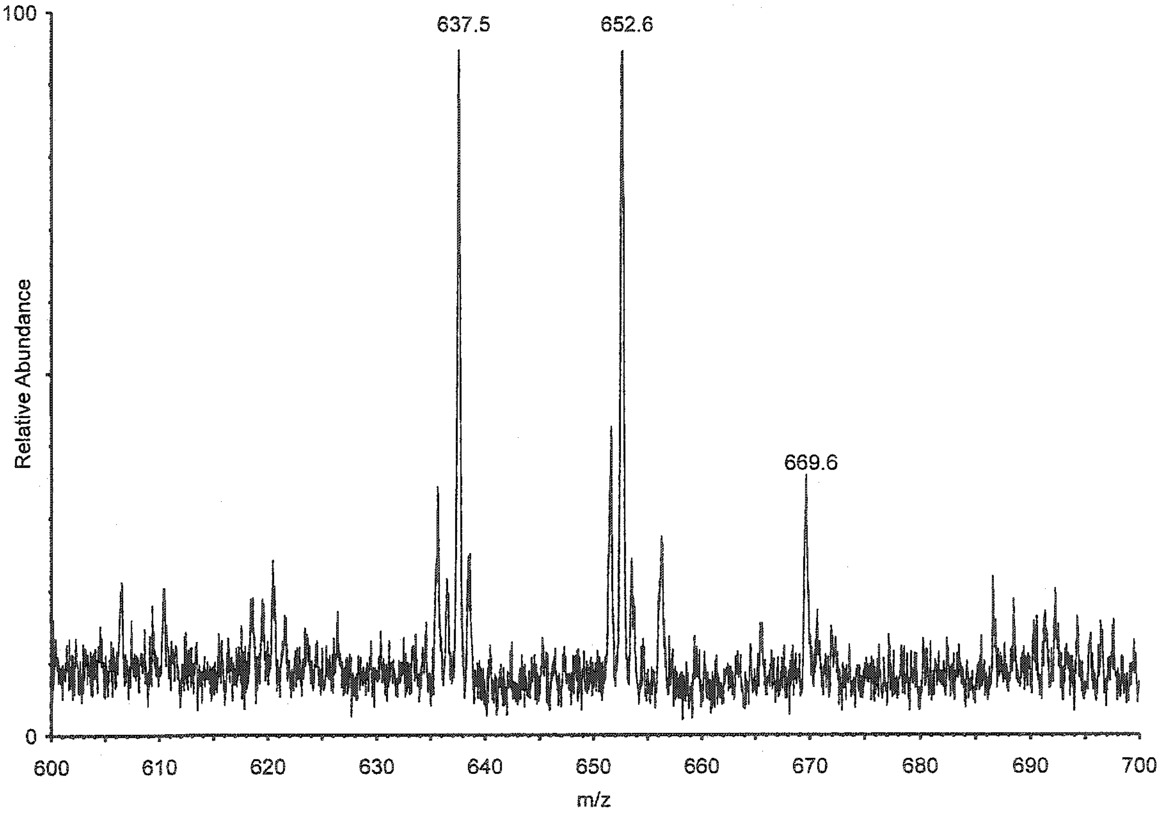


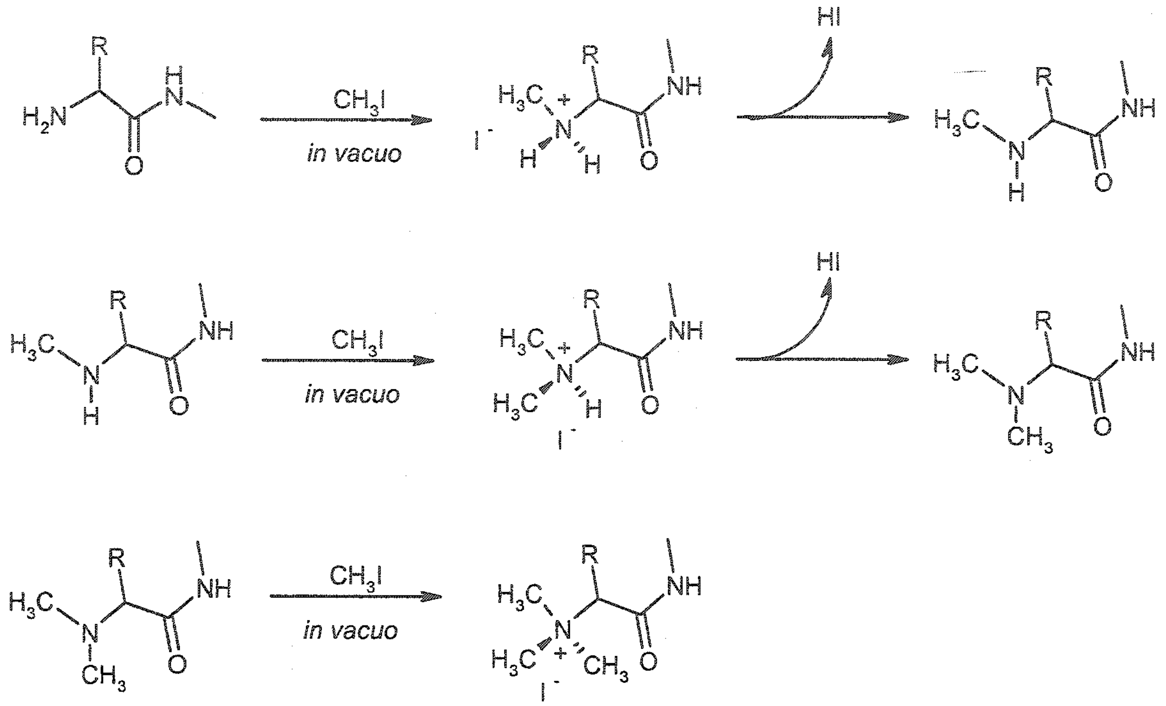
Figure 1-28 Portion of the MALDI MS displaying the signals corresponding to the N-terminal peptide $(CH_3)_3N^+$-VH(*imN*$_1$(CH_3),*imN*$_3^+$(CH_3))LTP-OH derived from the papain digest of an equal mixture of CH_3 and CD_3 *in vacuo* methylated hHb

An equivalent peptide derived from the N-terminal of the α-chain of hemoglobin was not found (N-terminal sequence starts with VLSPADKTN…). This could be due to the generation of peptide fragments which are too small (m/z < 200) to be observed under

MALDI as the matrix itself interferes in this mass range. To avoid this, a digest using a different enzyme or chemical cleavage could be tried to produce peptides which are long enough from the N-terminal of both chains of hHb.

In vacuo methylation of a peptide lyophilized from acids

Although a small amount of the mono and dimethylammonium derivative of amino groups are present and can be observed by NMR for proteins (Taralp and Kaplan, 1997) and by ESI MS for peptides (data not shown) the predominant derivative formed by *in vacuo* methylation is the trimethylammonium derivative. One may expect that based upon the "pH memory" effect, in that there are no protonic equilibria, the methylation would stop at the first step with the production of the monomethylammonium iodide derivative without the use of a proton scavenger such as a hindered base to "mop up" the acid produced during the reaction. Since the methylations were carried out under conditions where a substantial portion of the amino group was protonated, this suggests that the protonated form of the amino group can undergo methylation under *in vacuo* conditions. A proposed mechanism is shown in Scheme 1-5 where the amino group is fully methylated to the trimethylammonium derivative with very little mono or dimethylammonium derivative produced. Therefore a possible mechanism is explained in terms of the vacuum removing the proton by the loss of HI gas producing a free amine which, in turn, can undergo further methylation with the subsequent loss of HI to the vacuum until the fully methylated derivative is obtained.



Scheme 1-5 Proposed mechanism for the *in vacuo* methylation of peptides and proteins; reaction is driven with the sequential loss of HI to the vacuum

Based on this hypothesis an experiment was attempted in which a peptide was lyophilized under acidic conditions and methylated *in vacuo* to see if it would be possible to obtain the trimethylammonium derivative from the completely protonated form of the amino group. Different acids were also tested to see if the counter ion would have an effect on the reaction. Figure 1-29 shows the MALDI MS of the *in vacuo* methylated test peptide H-SQNYPIV-OH lyophilized from formic acid. It is clear that the reaction is not complete as the signals at 820.5 m/z for $[M + H]^+$ and 842.5 m/z for $[M + Na]^+$ of the underivatized peptide are apparent but the predominant signal is the peak at 876.6 m/z which corresponds to the peptide possessing 4 methyl groups.

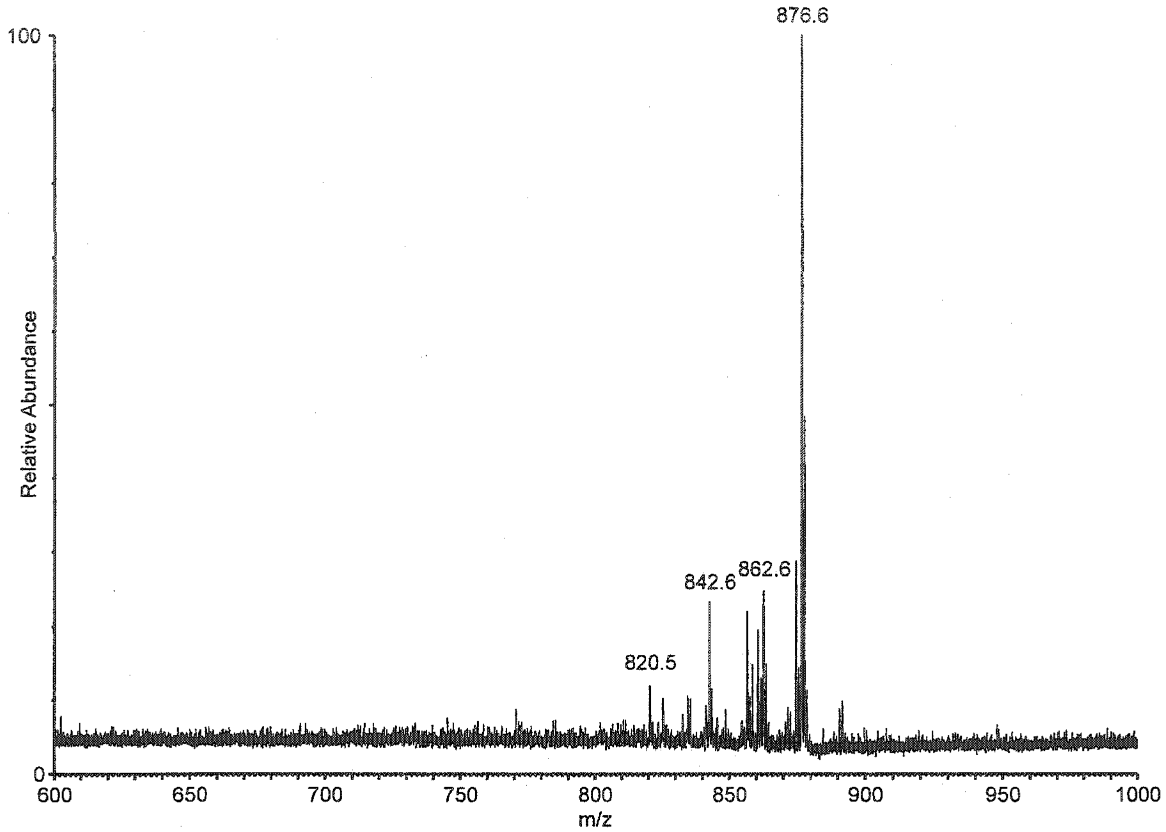


Figure 1-29 MALDI MS of the *in vacuo* methylated peptide H-SQNYPIV-OH lyophilized from formic acid

Similar results were obtained with the other acids; acetic acid, boric acid, HCℓ and TCA (see Figure 1-30).

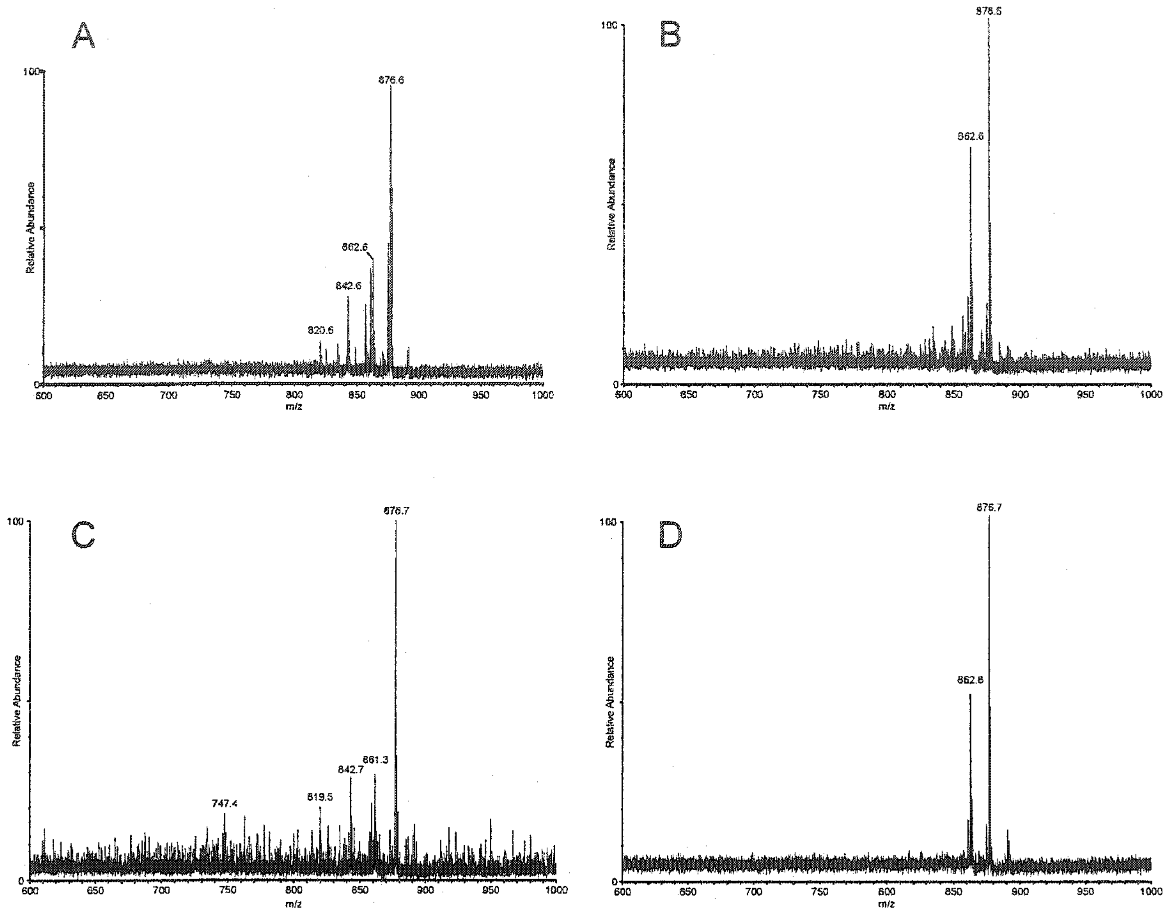


Figure 1-30 MALDI MS of the *in vacuo* methylated peptide H-SQNYPIV-OH lyophilized from A) acetic acid, B) boric acid, C) HCℓ and D) TCA

To compare the extent of methylation of the peptide lyophilized from an acid vs. the same peptide lyophilized under basic conditions, the MALDI mass spectrum of an equal mixture of *in vacuo* CH_3-methylated H-SQNYPIV-OH lyophilized from TFA and CD_3-methylated H-SQNYPIV-OH lyophilized from pH 8.0 buffer was recorded. As can be seen the same extent of methylation is observed (see Figure 1-31). A pair of predominant peaks, separated by 12 mu, is observed at 876.6 and 888.7 m/z, corresponding to $N^+(CH_3)_3$-SQNYPIV-OCH_3 and $N^+(CD_3)_3$-SQNYPIV-OCD_3, respectively.

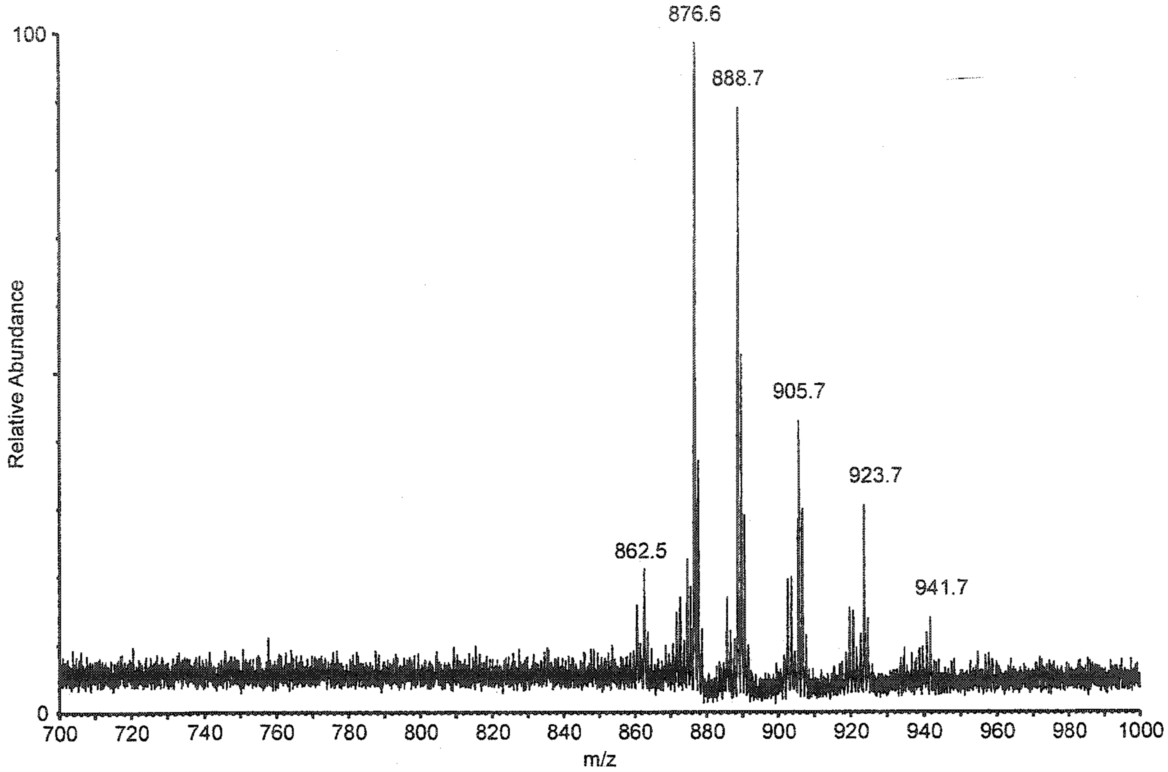
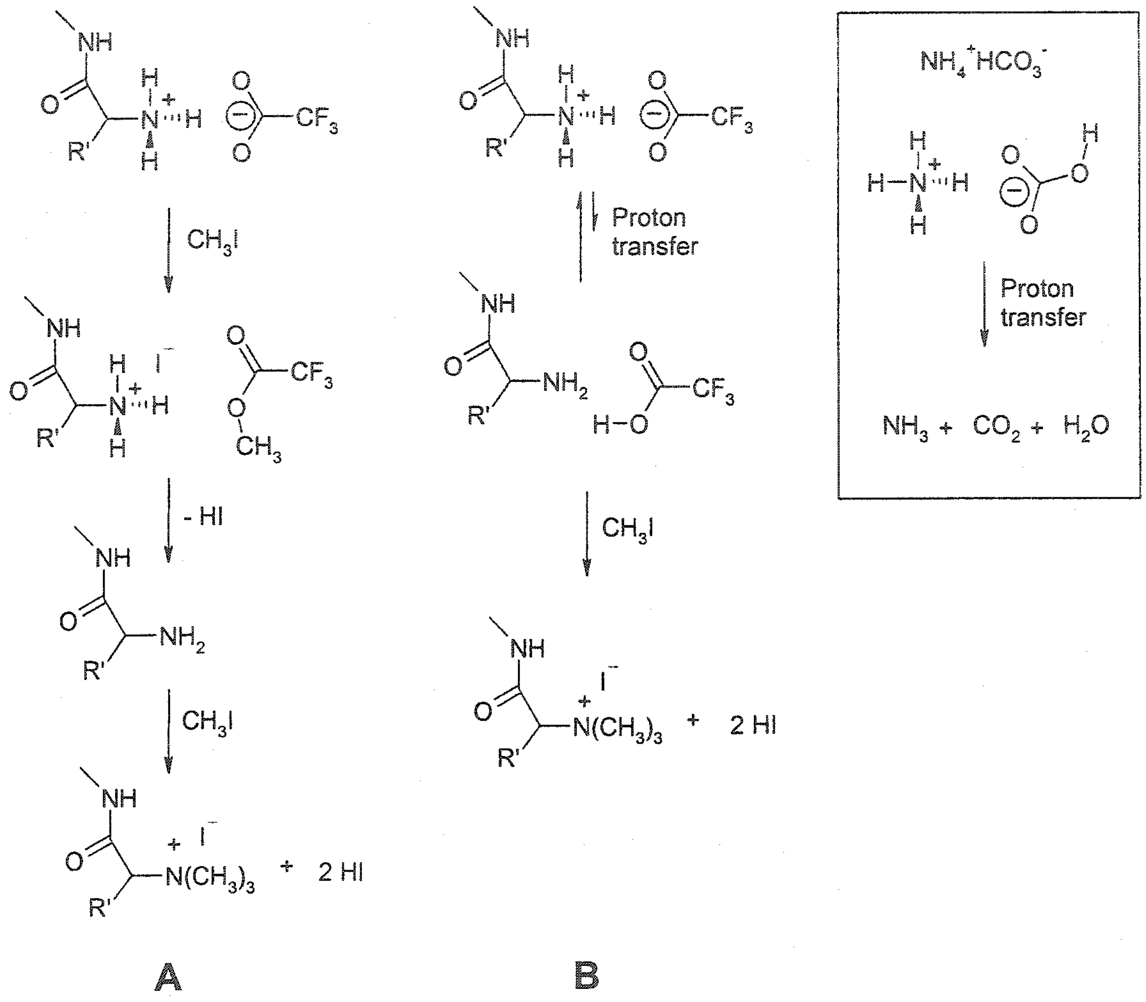


Figure 1-31 MALDI MS of an equal mixture of *in vacuo* CH_3-methylated H-SQNYPIV-OH lyophilized from TFA and CD_3-methylated H-SQNYPIV-OH lyophilized from pH 8.0 buffer

$N^+(CH_3)_3$-SQNYPIV-OCH_3 $[M]^+ = 876.6$ m/z, $N^+(CD_3)_3$-SQNYPIV-OCD_3 $[M]^+ = 888.7$ m/z

The fact that methylation of the peptide lyophilized from acids proceed to such a great extent leads to two possible mechanisms as proposed in Scheme 1-6. The first is the initial methylation of the counter-ion, TFA as example with proton transfer from the ammonium group to iodide to be removed as a gas to the vacuum. Trifluoroacetic acid methyl ester (TFAMe) has a boiling point of 43°C (Solvay Fluorides Inc Datasheet, Solvay Fluor & Derivate GmbH) which would also be lost to the vacuum which may further permit the methylation of the amino group. The second mechanism has the proton transfer occurring first, with subsequent methylation of the free amine, essentially wiping out the "pH memory". This would be analogous to the sublimation of $NH_4^+HCO_3$

in vacuo to $NH_{3(g)}$, $H_2O_{(g)}$ and $CO_{2(g)}$. The next step of the reaction mechanism would be that as described in Scheme 1-5.



Scheme 1-6 Two proposed mechanisms for the *in vacuo* methylation of peptides starting from the protonated amino group with (for example) TFA as counter-ion: **A** Initial methylation of the counter ion (trifluoroacetate) with subsequent proton loss *via* $HI_{(g)}$ to the vacuum and, **B** Initial proton transfer to carboxylate followed by rapid methylation of free amine. An analogous proton transfer is shown with $NH_4^+HCO_3^-$ (inset)

Whatever the mechanism, the fact is that the trimethylammonium is the predominant derivative formed by the *in vacuo* methylation of amino groups and this can occur from the completely protonated form of the amino group.

Finally, an argument could be made that the increased detection observed may be attributed, in part or in whole, to the ion suppression effect (Kratzer *et al.*, 1998; Knochenmuss *et al.*, 2000). The trimethylated peptide ion will carry the majority of the current, effectively suppressing all other signals. Perhaps the introduction of the TMA phosphate buffer prior to lyophilization may also have a suppression effect. Regardless of the cause, an increase in detection is observed with trimethylation of the amino group. The effect of trimethylation in peptides possessing two or more amino groups has not been investigated. It is difficult to predict what effect two permanent positive charges on a peptide may have during MALDI. Singly charged ions have been described as the "lucky survivors" in MALDI MS (Karas *et al.*, 2000) and as a result, the doubly charged ions may be suppressed under MALDI.

The increase in signal intensity for the trimethylammonium derivatized peptides has not been quantified. However, in each case, the trimethylammonium derivatized peptide had a greater signal intensity than the equivalent non-derivatized peptide. The increase in this sensitivity for the trimethylammonium derivatized peptide *vs*. the non-derivatized peptide could be determined by the slope of the relative signal intensity *vs*. the amount of peptide. This increase in sensitivity is predicted to be peptide dependent.

Chapter 2 Application of *in vacuo* methylation for the isolation of proteins' C-terminal peptides

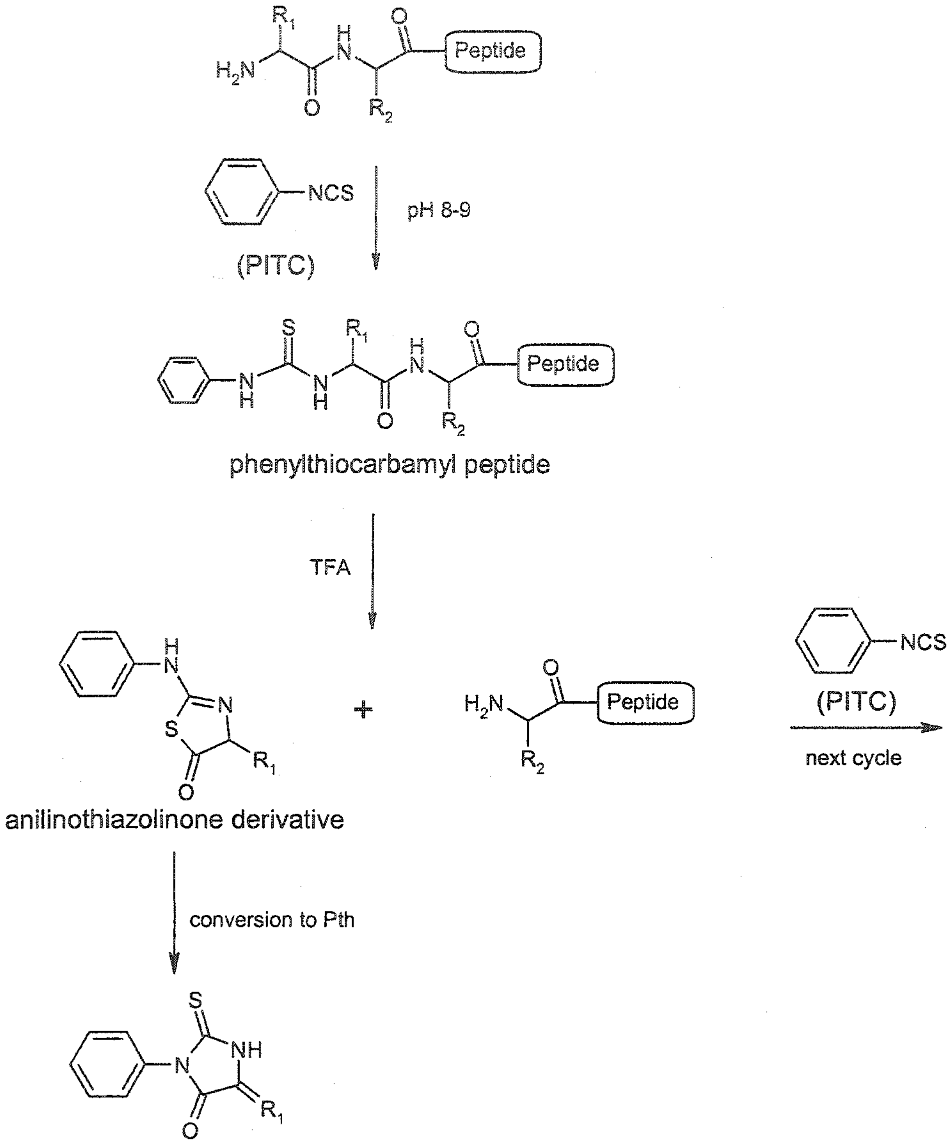
Introduction

Structure-function relationship investigations of proteins can be greatly facilitated with the knowledge of the various levels of protein structure. The primary structure, *i.e.* the amino acid sequence, is the foundation in the architecture of higher levels of protein structure. Much progress has been made in determining the primary structure of proteins by the introduction of deoxyribonucleic acid (DNA) sequencing, where the protein's sequence is obtained indirectly by translation of its DNA sequence for which it encodes. Before the development of DNA sequencers, the main strategy employed to determine the full amino acid sequence of a protein was by the N-terminal Edman sequential chemical degradation of individual peptide fragments produced by chemical cleavage (*e.g.* CNBr) and different enzymatic digestions (*e.g.* trypsin, V8 protease), followed by final alignment of overlapping sequences to provide the complete sequence. The assignment of such sequences is often ambiguous due to missing sequences and is labour intensive therefore this approach is not ideal (Creighton, 1993). Although its use for complete protein sequencing has declined, the Edman degradation chemistry (Edman and Begg, 1967) is still used to characterize the N-terminus of synthetic peptides, recombinant proteins, post-translations proteolytic processing and as an initial characterization of purified proteins by matching the amino acid sequence obtained to the DNA sequence. The degradation chemistry (see Scheme 2-1) is a cyclic procedure, where amino acids are cleaved as their anilinothiazolinone derivative with an anhydrous acid one at a time from the N-terminus of the peptide after reaction with phenyl isothiocyanate (PITC) to form a phenylthiocarbamyl peptide. The anilinothiazolinone derivative is converted to a more stable phenylthiohydantoin (Pth) derivative for identification following chromatographic separation. Since its original development, the methodology can be carried out on 5 to 10 picomoles, and sequences in the range of 30 to 70 amino acids can be determined on an automated sequencer (Hunkapiller and Hood, 1983) provided the peptide or protein sample is relatively pure of other peptide or protein impurities and, for a protein, it must consist of one polypeptide chain with a free α-amino group. Blocked N-terminal groups *e.g.* formyl (f-Met), acetyl, and pyroglutamate residues formed by cyclization of N-terminal glutamine or glutamate, are common and present a formidable shortcoming to this chemistry.

N-terminal sequencing of proteins by this cyclic degradative chemistry has proven to be easier than sequencing from the C-terminus, simply due to the higher nucleophilic character of the α-amino group of proteins and peptides *vs.* the α-carboxylate group. Despite many attempts there is no satisfactory method similar, in sensitivity and in length of passage, to the Edman for C-terminal sequencing. An isothiocyanate (ITC) method (Schlack and Kumpf, 1926) similar to that of the Edman has been investigated as a potential cyclic sequencing method for automated applications. The method consists of three steps: activation of the α-carboxylate to a mixed anhydride, derivatization to the peptidylisothiocyanate with an ITC reagent, followed by cyclization to the peptidylthiohydantoin and hydrolysis of the derivatized C-terminal amino acid thiohydantoin to an amino acid hydantoin and a shorten peptide (see Scheme 2-2). Many types of reagents have been tried as PITC employed in the Edman degradation is not nucleophilic enough for the C-terminal method. Both the Edman and the modified

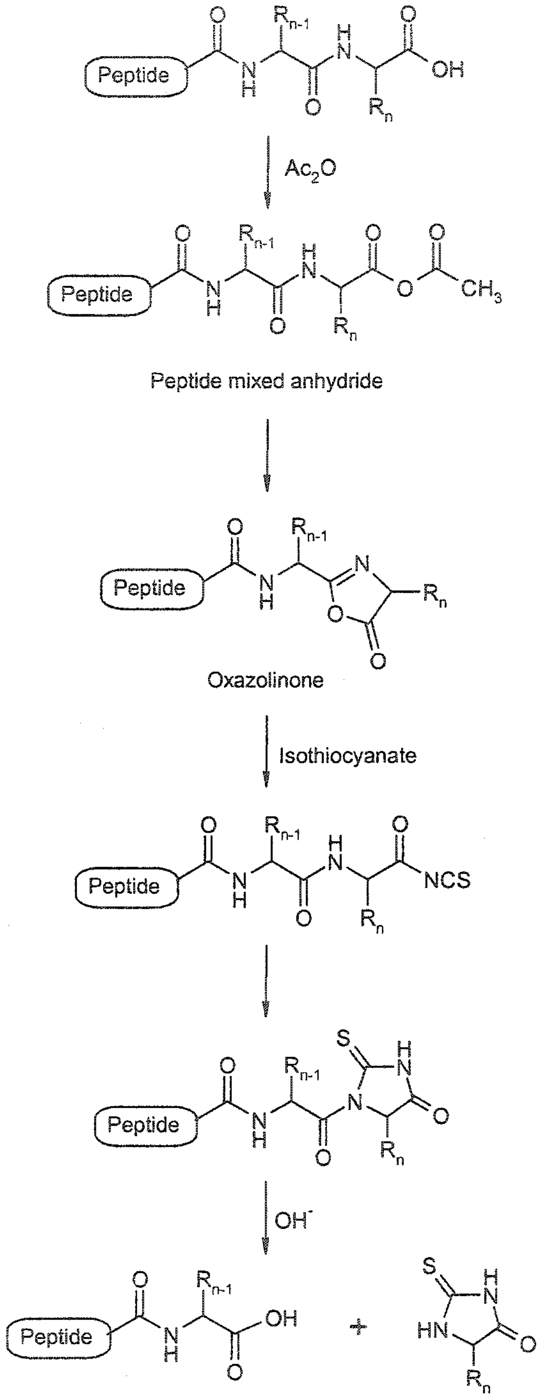
Schlack, Kumpf methods rely on a time consuming chromatography separation on a dedicated instrument for the identification of the amino acid hydantoins. Despite many improvements in the chemistry, the amino acid proline, which contains a secondary α-amino group, continues to be problematic as the hydantoin derivative produced during the C-terminal sequencing method is not stable. Proline passage has been reported (Bergman *et al.*, 2001) but it is not observed in the chromatographic trace and can only be inferred to by the absence of an amino acid hydantoin signal in the cycle. Because the chemistry is not completely quantitative, each cycle will possess a signal in the trace from the predecessing amino acids in the peptide sequence which severely complicates sequence interpretation in the case of repeating amino acids and with the presence of proline.

The ability to obtain the C-terminal amino acid sequence of proteins is of great interest for: (1) confirmation of the C-terminal sequence of natural and recombinant proteins or peptides, (2) identification of post translational modification and proteolytic processing at the C-terminus, (3) sequencing of N-terminally blocked peptides or proteins, (4) to produce oligonucleotide probes for cloning cDNAs, and (5) to confirm sequence data obtained indirectly from the DNA sequence.



Scheme 2-1 N-terminal sequencing of peptides and proteins by cyclic Edman degradation

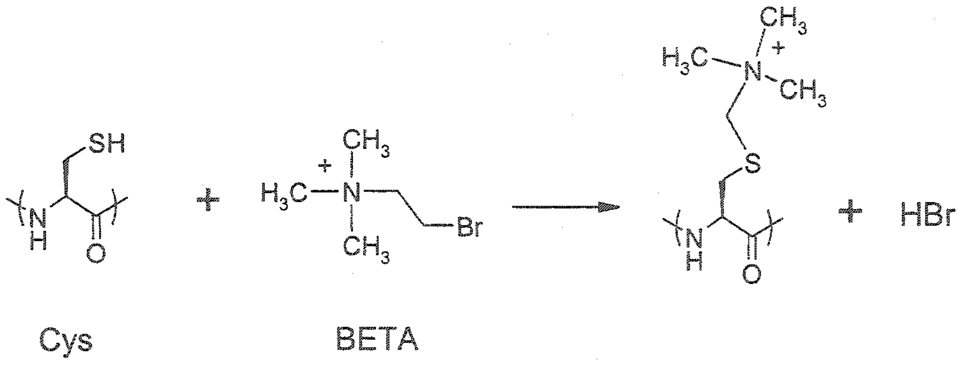
The peptide is first reacted with PITC under basic conditions. The phenylthiocarbamyl peptide is cleaved under acid conditions producing an anilinothiazolinone derivative, which is converted to a stable Pth and the (n-1) peptide which can proceed to the next cycle.



Scheme 2-2 C-terminal degradation of peptides

Therefore a C-terminal sequencing method with general applicability is greatly needed. This need has initiated a number of other strategies for the determination of

C-terminal sequence of proteins and peptides. One approach makes use of carboxypeptidases to cleave amino acids, ideally one at a time, from the C-terminus. However, these enzymes have preferred substrate specificities and display different cleavage kinetics, therefore a mixture of carboxypeptidases A, B and Y are often used to compensate for this (Tsugita, 1987). In theory, the C-terminal amino acids could be analysed as they are released (*i.e.* amino acid analysis) but the strategy employed is to analyse the resulting truncated peptides produced after periods of time by MS and the differences in their masses measured to deduce the sequence (Tsugita, 1987; Thiede *et al.*, 1995). This method can be problematic as it cannot distinguish between the isobaric amino acids leucine and isoleucine, and the amino acids lysine and glutamine which also have very similar masses and therefore require MS instruments of high accuracy. Chemical modification of lysine (*e.g.* by guanidinylation, see Scheme 1-1) to differentiate this amino acid from glutamine has been attempted with some success using MALDI MS (Bonetto *et al.*, 1997) but again this methodology requires some amount of purification prior to MS. Cysteine containing peptides pose another problem in this strategy for the fact that of carboxypeptidase (CP) Y and CP P do not cleave peptides that contain non-derivatized cystine, cysteic acid (from performic acid oxidation) or (carboxymethyl)cysteine. Again, chemical modification has been attempted by converting cysteine residues to 4-thialaminine; a positively charged derivative, by (trimethylamino)-ethylation with (2-bromoethyl)trimethylammonium bromide (BETA), see Scheme 2-3 (Bonetto *et al.*, 1997; 1997).



Scheme 2-3 Derivatization of Cys residues with BETA to 4-thialaminine for passage by CP Y and P (from (Bonetto *et al.*, 1997; 1997))

It is interesting to note that 4-thialaminine, with a permanent positive charge, is very similar in structure to the trimethylammonium derivative of amino groups produced by the *in vacuo* methylation of peptides and proteins (*vide supra*) but an increase in detection in MALDI MS was not mentioned (Bonetto *et al.*, 1997; 1997).

As well as the methods mentioned above there exists one more strategy. That is to somehow separate out and isolate a peptide derived from the C-terminus of a protein for further characterization. One technique employs anhydrotrypsin, a catalytically inert derivative of trypsin that binds peptides containing lysine or arginine residues at their C-terminal. A tryptic digest of the protein of interest is passed over a column containing immobilized anhydrotrypsin. The peptides are retained on the column except for the C-terminal peptide provided it does not possess a C-terminal arginine or lysine. This would be observed as the absence of a peptide eluting from the column and consequently the purity of the protein of interest is of the highest concern. If the C-terminal peptide does possess a C-terminal arginine or lysine this shortcoming can be circumvented by applying the digest from an enzyme of different specificity other than trypsin. Another limitation to this method is that the peptides produced from a tryptic digest are often small as arginine and lysine are quite common, therefore insufficient sequence information is gained. Again, attempts have been made to overcome this by using the endoprotease Lys-C, which produces peptides with C-terminal lysines. The C-terminal peptide in this case was obtained by comparing MALDI MS spectra of the digest prior to and after incubation with anhydrotrypsin (Sechi and Chait, 2000). A disadvantage to this strategy is that any trace endoproteases will generate peptides leading to false C-terminal peptides therefore the enzyme preparation used must be extremely pure.

Another approach is to chemically modify the carboxyl groups of a protein and to isolate the C-terminal peptide based on this new chemical property (Furka *et al.*, 1970; Duggleby and Kaplan, 1975; Furka *et al.*, 1983). This strategy was explored using the *in vacuo* methylation procedure to specifically modify the carboxyl groups of proteins by converting them to their methyl esters.

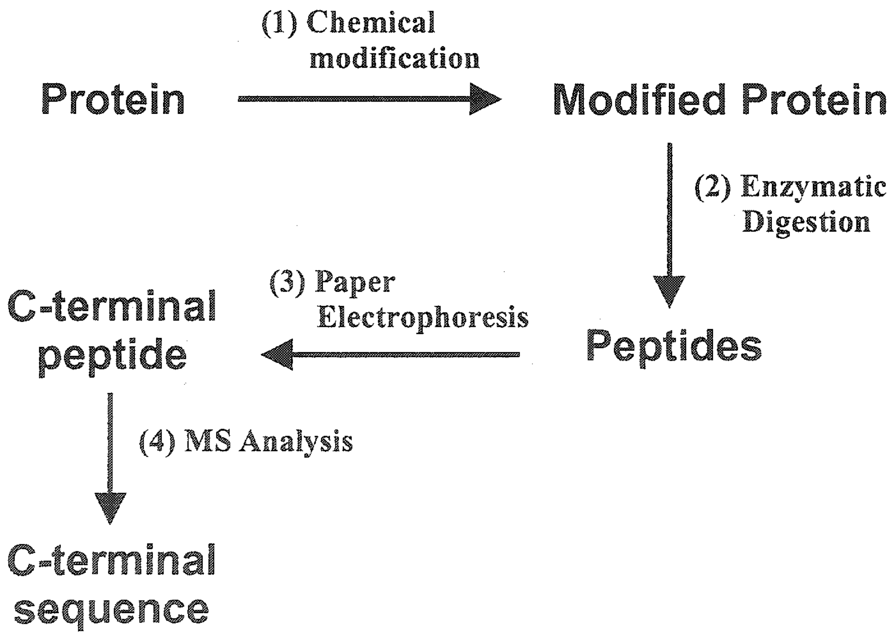
Objectives

The first objective was to determine if the *in vacuo* methylation procedure of proteins could be effectively applied to differentially label the C-terminal end of proteins for further isolation of the C-terminal peptides from a proteolytic digest by an electrophoretic separation. The second objective was to determine the most favourable conditions for the proteolytic digestions of the methylated protein to produce peptides of long enough length. The ultimate objective is to develop a method of general applicability.

Strategy for the isolation of C-terminal peptides

The procedure explored and developed is a variation of the C-terminal methodology employing "diagonal" electrophoresis.

The novelty of the proposed approach employed for the isolation of C-terminal peptides involves the *in vacuo* chemical modification of the C-terminal carboxyl groups of the protein, followed by enzymatic cleavage of the polypeptide, and finally, isolation and sequence determination of the carboxyl methylated terminal peptides, as illustrated in Scheme 2-4.



Scheme 2-4 Strategy employed to isolate C-terminal peptides of proteins for potential sequence

In the chemical modification step (Scheme 2-4, step1), the carboxyl groups of the proteins are converted to their methyl esters by *in vacuo* methylation with iodomethane.

An equal mixture of CH_3 and CD_3 iodomethane is used for potential MS analysis. The protein is first methylated *in vacuo* with a trace amount of ^{14}C-iodomethane for visualization purposes by autoradiography. Methylation of the protein can be controlled to a certain extent, to target predominantly the carboxyl groups, by simply adjusting the pH. Enzymatic cleavage of the methylated protein produces peptide fragments (Figure 2-1, step 2) with a free carboxy terminal except for the original C-terminal peptide. The pK_a values of the terminal carboxyl groups of proteins range from 4 to 5, whereas the pK_a values of the free amines range from 7 to 9 (Creighton, 1993). Therefore at pH 2.1 (pH for the volatile buffer used), the free carboxyl groups of peptides are protonated, whereas at pH 4.4 (pH of the second dimension buffer), these groups will be deprotonated (see Figure 2-2), resulting in the gain of one negative charge in the peptide in going from pH 2.1 to pH 4.4 (Table 1). However, this is not the case for peptides derived from the C-terminus due to the fact that their C-terminal carboxylate groups have been esterified. Two-dimensional high voltage paper electrophoresis (2D HVPE) separation of the peptides, with the first dimension run at pH 2.1 and the second at pH 4.4, will place the C-terminal peptides on a diagonal of the two dimensions, as illustrated in Figure 2-3. This is due to the fact that their electrophoretic mobility will remain the same under pH 2.1 and pH 4.4, whereas the other peptides will gain a negative charge and will not travel as far in the second dimension and as a result will lie below the diagonal.

In theory, mass spectrometry can be used to determine the sequence of the isolated C-terminal peptides by MS/MS. Thus, longer C-terminal peptides isolated by 2D HVPE will provide the greatest sequence coverage. This is particularly useful since the knowledge of longer sequence significantly increases the probability of identifying the protein based on existing libraries or databases. The identification of C-terminal peptides from non-peptide impurities can be greatly simplified by the use of mixtures of isotopically labelled (CH_3 and CD_3) iodomethane.

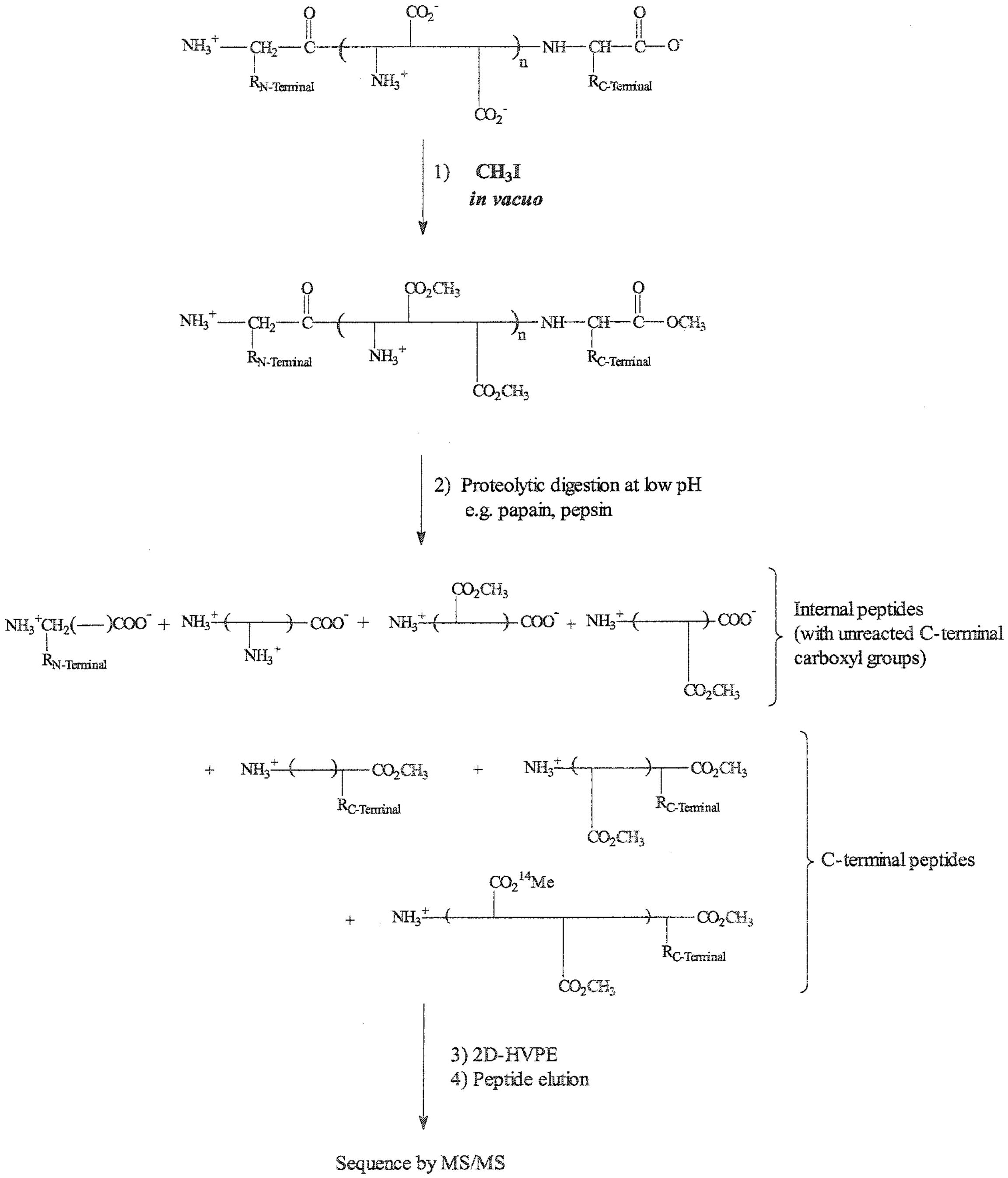


Figure 2-1 Strategy for the determination of the C-terminal sequence of proteins

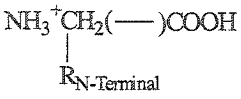
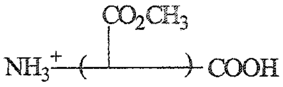
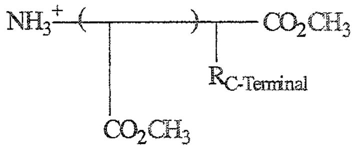
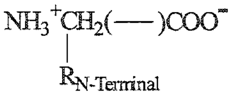
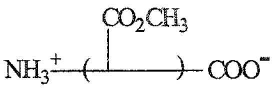
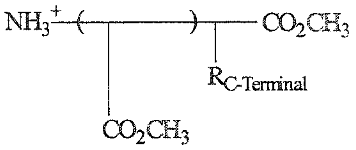
pH	N-terminal and internal peptides	Original C-terminal peptides
pH 2.1	 $\mathrm{NH_3^+\underset{R_{N\text{-}Terminal}}{CH_2}(—)COOH}$  $\mathrm{NH_3^+\text{—}(\overset{CO_2CH_3}{\text{—}}\text{—})\text{—}COOH}$	 $\mathrm{NH_3^+\text{—}(\underset{CO_2CH_3}{\text{—}}\text{—})\underset{R_{C\text{-}Terminal}}{\text{—}}CO_2CH_3}$
pH 4.4	 $\mathrm{NH_3^+\underset{R_{N\text{-}Terminal}}{CH_2}(—)COO^-}$  $\mathrm{NH_3^+\text{—}(\overset{CO_2CH_3}{\text{—}}\text{—})\text{—}COO^-}$	 $\mathrm{NH_3^+\text{—}(\underset{CO_2CH_3}{\text{—}}\text{—})\underset{R_{C\text{-}Terminal}}{\text{—}}CO_2CH_3}$

Figure 2-2 Comparison of N-terminal, internal and C-terminal peptides at pH 2.1 and pH 4.4

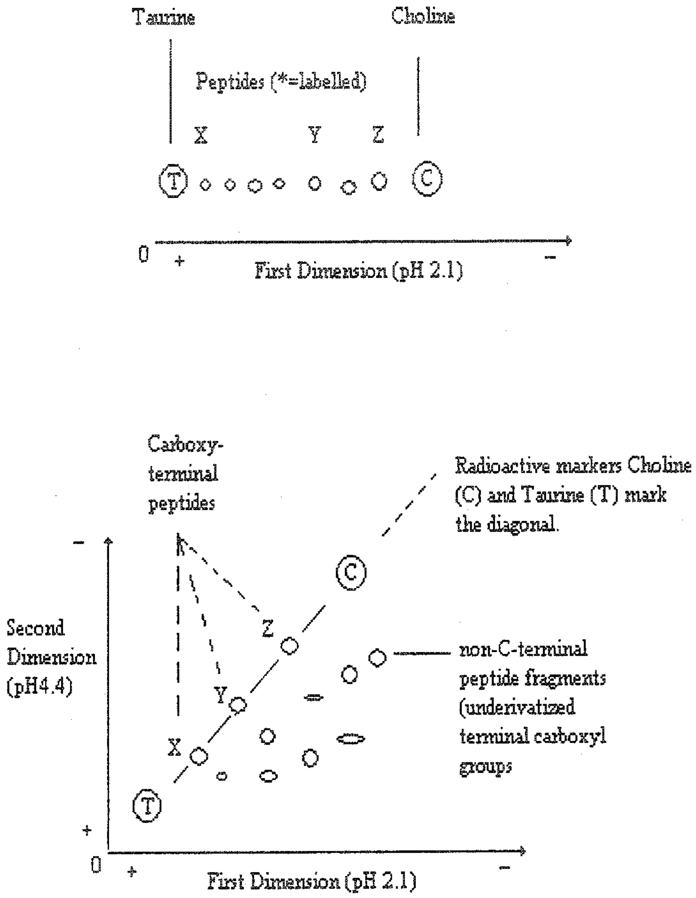


Figure 2-3 Isolation of C-terminal peptides using two-dimensional high voltage paper electrophoresis (2D-HVPE)

Materials and Methods

αCT, Type II from bovine pancreas (54 units/mg protein, Lot#27H7007), pepsin from porcine stomach mucosa (2750 units/mg protein, Lot#128C8112), papain from papaya latex (suspension in 0.05 M sodium acetate, 24 units/mg protein, Lot#124F8045), carboxypeptidase Y (CP Y) from *Saccharomyces cerevisiae* (Lot#106H8060), carboxypeptidase A (CP A) from bovine pancreas (Lot#88H7425), carboxypeptidase B (CP B) from porcine pancreas (Lot#75F8046) and pronase E (from *Steptomyces griseus,* purified type XIV, Lot#118F0732) were purchased from Sigma. All enzymes except pronase were used without further purification. Pronase (100 mg in 30 ml ddH_2O, pH adjusted to 4.0 with 0.1 M HCl) was dialysed (MWCO 3,500) against 1 % acetic acid (2 X 3.5 L), and lyophilized.

The tripeptide H-Thr-Tyr-Ser-OH (Lot#88H5821) was obtained from Sigma. The hexapeptide H-Ser-Ala-Ser-Ala-Phe-Ala-OH, synthetic, was gratefully provided by C.T. Choma, RPI Troy NY, purified by RP HPLC, prior to use

$^{14}CH_3I$ (specific activity 55.0 mCi/mmol) was purchased from Amersham Chemicals and $^{14}CH_3OH$ was obtained from NEN 250 μCi. CD_3I was obtained from Cambridge Isotope Laboratories, Inc. and CH_3I was purchased from Aldrich. 7-Methoxycoumarin-3-carboxylic acid, succinimidyl ester was purchased from Molecular Probes. Spermidine and ethanolamine were obtained from Aldrich. X-ray film for autoradiography was from Fuji. All other reagents and solvents were of high-purity obtained from commercial sources.

The following volatile electrophoresis buffers were prepared:

pH 2.1 buffer: glacial acetic acid/ formic acid/ dH_2O (4:1:45)

pH 4.4 buffer: glacial acetic acid/ pyridine/ dH_2O (4:3:600)

pH 6.5 buffer: glacial acetic acid/ pyridine/ dH_2O (3:100:900)

The following stock solutions of fluorescent markers were prepared:

5-dimethylamino-1-naphthalenesulfonyl arginine (Dns-Arg) 0.2 mg/mL in pH 2.1 volatile buffer and 5-dimethylamino-1-naphthalenesulfonic acid (Dns-Sulfonic) 0.1 mg/mL in pH 6.5 volatile buffer.

The following radiolabeled markers were prepared:

^{14}C-Taurine ($H_2N^{14}CH_2{}^{14}CH_2SO_3H$; 2-aminoethanesulfonic acid): 5 μCi/mL in pH 2.1 buffer and ^{14}C-choline ($HOCH_2CH_2N(^{14}CH_3)_3 \cdot Cl$): 5 μCi/mL pH 2.1 buffer.

Preparation of lyophilized αCT at LpH 4.5

αCT (206 mg) was dissolved in dH_2O (16 mL) to yield a concentration of approx. 13 mg/mL. The pH of the solution was adjusted to pH 4.5 using 1.0 M NaOH delivered *via* syringe, where the final pH was monitored with a semi-micro gel-filled combination electrode (Orion) pH meter (Model PHM26; Radiometer Copenhagen). Equal aliquots (2.0 mL) of the proteins were withdrawn and transferred to individual disposable borosilicate glass culture tubes (x 8 tubes) and the tubes were immediately frozen in liquid nitrogen. Samples were lyophilized overnight.

In vacuo trace methylation of αCT with ^{14}C-iodomethane

The procedure for *in vacuo* radiolabelling of proteins is summarized in Figure 0-1. A hydrolysis tube containing the lyophilized protein sample (ca. 20 mg) was elongated and narrowed with an opening of approx. 0.2-0.5 mm diameter by heating and stretching it over a hot flame. ^{14}C-Iodomethane in octane (0.5 mCi/mL; 60 μL) was added *via* pipette in the narrowing of the tube, the bottom of the tube was immediately immersed in liquid nitrogen, creating a suction pulling the reagent into the tube and trapping it on the sides preventing it from "wetting" the protein. The tube was placed under vacuum, flame-sealed and allowed to react in an oven preset at 75°C for 24 hours. The reaction was stopped by trapping out the $^{14}CH_3I$/octane at the top of the reaction vessel with liquid nitrogen. Scoring the glass tube and applying a heated glass rod to it, cracking the tube, subsequently released the vacuum. The sample was re-labelled with $^{12}CH_3I$ following the same procedure as above for homogeneity.

In vacuo double isotope labelling of αCT

CH_3I:CD_3I (1:1 mol. equiv.; 30 μL) was added to each tube of αCT in the same manner as for radioactive labelling procedure. The tubes were frozen in liquid nitrogen, sealed under vacuum and allowed to react overnight in an oven at 75°C. The reaction was terminated as above.

Preparation of pepsin and papain enzymatic stock solutions

Pepsin (2.3 mg) was dissolved in freshly made 10% formic acid (5.0 mL) to give a pepsin solution with concentration of 0.4 mg/mL. Papain (40 μL) was dissolved in pH 4.5 buffer (950 μL), and activated with β-mercaptoethanol (BME) (10 μL) to give a final concentration of enzyme of 1 mg/mL.

Enzymatic digest of ^{14}C-methylated αCT

Pepsin digest

The ^{14}C-methylated methylated protein samples (ca. 5 mg) were transferred to glass scintillation vials and dissolved in formic acid (98%; 500 μL) and then made up to 10% formic solution with distilled water (4.5 mL). These were each divided equally into 5 aliquots of 1 mL each. Pepsin enzymatic solution (25.0 μL) was added to each aliquot (noting time zero) to give a protein/enzyme ratio of 100:1. The enzymatic digestion was performed in a shaker water bath at 37°C for defined time intervals of 15 minutes, 30 minutes, 1 hour, 2 hours and 24 hours. Once the digest was stopped, the peptides generated were dissolved in dH_2O (5 mL), frozen over liquid nitrogen and lyophilized.

Papain digest

The ^{14}C-methylated protein samples (ca. 5 mg) were transferred to glass scintillation vials by dissolving them in formic acid (50 μL), followed by three washings (100 μL) of ddH_2O and totalled to 500 μL (10% formic acid) with further addition of ddH_2O. The samples were made up to a final volume of 5.0 mL with pH6.5 volatile buffer, pH was verified with pH strip, and divided equally into 5 aliquots of 1 mL each (1 mg protein/mL). A papain stock solution was prepared (1.0 mg/mL in pH 6.5 volatile buffer with 1% (v/v) BME). An aliquot (20 μL) of the papain stock solution was added to each aliquot (noting time zero) to give a protein/enzyme ratio of 50:1. The enzymatic digestion was incubated in a water bath at 37°C and allowed to proceed at the time intervals of 15 minutes, 30 minutes, 1 hour, 2 hours and 24 hours. The reactions were stopped and lyophilized as above.

Two-dimensional high voltage paper electrophoresis

Apparatus

The apparatus employed for high-voltage paper electrophoresis (HVPE) consisted of a 40 L glass chromatography solvent tank containing an insulating/cooling solvent (Varsol™, ExxonMobil Chemical, for pH 2.1 and pH 4.4 electrophoresis systems and toluene for the pH 6.5 system). The tank was equipped with a lid fitted with glass cooling coils in which tap water would flow. The wet chromatography paper would hang over a glass rod with one end of the paper at the bottom of the tank immersed in a layer of buffer containing the anode, and the other end was placed atop in a trough filled with buffer containing the cathode. Once closed (confirmed by safety interlock switches) the system would be powered up and the current would flow through the paper.

Preparation of sample for first-dimension HVPE

The chromatography paper (Whatman 3MM, 40cm x 57cm) was placed on a clean glass surface with the origin line raised by two glass rods. The enzyme-digested αCT sample (ca. 5 mg) was dissolved in pH 2.1 buffer (500 μL). The solution of the dissolved sample was applied *via* capillary pipette along the origin line of the paper as a band 5 cm wide. Dns-Arg (20 μL) was applied at the edges and in the middle of the band as fluorescent marker. A mixture of ^{14}C-choline and ^{14}C-taurine (30 μL) marker was applied in the middle of the band in the same manner (to delineate the diagonal in the second dimension). The sample and markers were dried after each application with a stream of warm air. The paper was folded along the origin and the whole paper was wetted with pH 2.1 buffer, applied with a squeeze bottle carefully at equal distance on each side of the fold, so that the buffer front concentrates the sample evenly along the origin line; a process known as "buffering-in". Excess solvent was removed from the paper by blotting until the paper was damp. The first dimension electrophoresis was performed for 30 min at 2.5 kV (voltage gradient of 40V/cm) in the tank containing pH 2.1 electrode buffer, at approximately 25°C. The paper was then removed from the tank with the aids of glass rods, and hanged to dry in a well-ventilated fumehood. When dried, viewing the fluorescent dye marker Dns-Arg previously spotted on the paper under UV light revealed the qualitative mobility of the sample band in the first dimension.

Second-dimension HVPE

With the aid of the fluorescent markers (Dns-Arg and Dns-Sulfonic), the sample band was cut out as a strip of paper and subsequently sewn onto a second full size chromatography paper aligning the strip along the origin of the new paper. The paper behind the strip was carefully removed with a razor blade to allow the current to flow through the sewn strip. The paper was then buffered-in with pH 4.4, in the same procedure as described above. Electrophoresis was again performed for 30 min at 2.5 kV (voltage gradient 40 V/cm) in the tank containing pH 4.4 electrode buffer. The paper was removed and treated in the same manner as above.

Autoradiography

The dried papers were marked with radioactive ink (prepared by the addition of waste ^{14}C isotopes to ink, ca. 1000 cpm/µL) to aid in the re-alignment of the developed film, and autoradiographed by direct contact to the film for at least 48 hours. The films were developed and fixed with Kodak fixing reagent, with a final rinse in distilled water before drying.

Peptide recovery

The diagonal line delineated by ^{14}C-taurine (origin) and ^{14}C-choline (extremity) on the paper electrophoretogram, containing the radioactive C-terminal peptides, was cut out. The cut out strips were eluted first with Varsol to remove any Varsol-soluble contaminants from the tank, then with pH 2.1 buffer to recover the C-terminal peptides. The buffer was subsequently removed by lyophilization.

Small HVPE apparatus

A small HVPE apparatus was designed using an Owl EasyCast™ horizontal minigel electrophoresis system (see Figure 2-8) placed in an ice bath. 3MM Whatman paper was cut to size (see Figure 2-9) and the wet paper was sandwiched between two sheets of polyethylene (15 cm x 10 cm) with each ends submerged in electrophoretic volatile buffer. A glass plate (10 cm x 10 cm) was placed on top followed by a Styrofoam block, the acrylic insert (provided with the system for the setting of an agarose gel) and a 20kg (9 in. x 9 in. x 1.5 in.) lead weight (carefully). Electrophoresis was performed at 200 V for 30 min.

Preparation of fluorescent markers for small HVPE

Fluorescent markers were prepared to delineate the diagonal instead of the radiolabeled markers previously used. Ethanolamine was chosen for the neutral marker and spermidine for a marker with a +2 charge after reacting them with 7-methoxycoumarin-3-carboxylic succinimidyl ester in 2% TEA for 1 hour. Markers were dried by Speed-Vac™, redissolved in pH 2.1 volatile buffer and quickly purified by HVPE by spotting some (200 μL) over a full sheet (minus border) and eluting off appropriate fluorescent band with pH 2.1 buffer to ca. 500 μL volume. Markers were stored in -20°C freezer.

Carboxypeptidase digest of the peptide: H-SASAFA-OH

Stock solutions for CP B and CP Y were prepared in NH_4HCO_3 (200 mM) pH 8.0, as these were obtained as dry lyophilized powders, whereas carboxypeptidase A was already in solution form. A stock mixture of CP A, B and Y was prepared by aliquoting appropriate amounts of each stock solution to obtain approximately equal amounts of activity for each enzyme based on the units per mg of protein given for each enzyme. Digestion of H-Ser-Ala-Ser-Ala-Phe-Ala-OH (10nmol) with pronase and the carboxypeptidases was carried out with the same carboxypeptidase stock as above with pronase added for a final of carboxypeptidase to pronase weight ratio of 50 to 1. The digestion was performed at ambient temperature (ca. 20°C) with an enzyme to peptide substrate ratio of 1 to 10. Aliquots were removed at timed intervals for up to 2 hours, combined to 50 µL 1% TFA to stop the digest and dried under a vacuum using a Speed-Vac™ apparatus.

In vacuo methylation of αCT with ^{14}C-MeOH·HCℓ

αCT (2 mg) was lyophilized from pH 2.1 buffer (2 mL) and methylated *in vacuo* as previously described (Vakos *et al.*, 2001) with ^{14}C-MeOH·HCℓ (10:1) 50 µL (250 µCi/ml methanol and 4M HCℓ in dioxane). The methylated protein was digested with pepsin as previously described (*vide supra*).

Results and Discussion

C-termini peptides of proteins derived from enzymatic digests were successfully separated and isolated by HVPE by the *in vacuo* reaction of lyophilized proteins with iodomethane. A typical example is shown in Figure 2-4. This represents a substantial improvement to the method previously described (Furka *et al.*, 1970; Duggleby and Kaplan, 1975; Furka *et al.*, 1983).

A typical example of an autoradiogram of a diagonal 2D-HVPE is shown in Figure 2-4. In this example, αCT (LpH 4.5) was trace methylated with ^{14}C-iodomethane followed by "cold" methylation and digested with papain (50:1 substrate: enzyme ratio) for 60 min. at 37°C. Radioactive markers taurine and choline are identified as spots at the origin and extremity of diagonal, respectively.

The large C-terminal peptide fragments have low electrophoretic mobility, and are indicated with solid arrows. The dashed line delineates the diagonal containing the C-terminal peptides. The 2D-HVPE for methylated αCT digested with pepsin and papain at various time intervals are shown in Figure 2-5 and Figure 2-6, respectively.

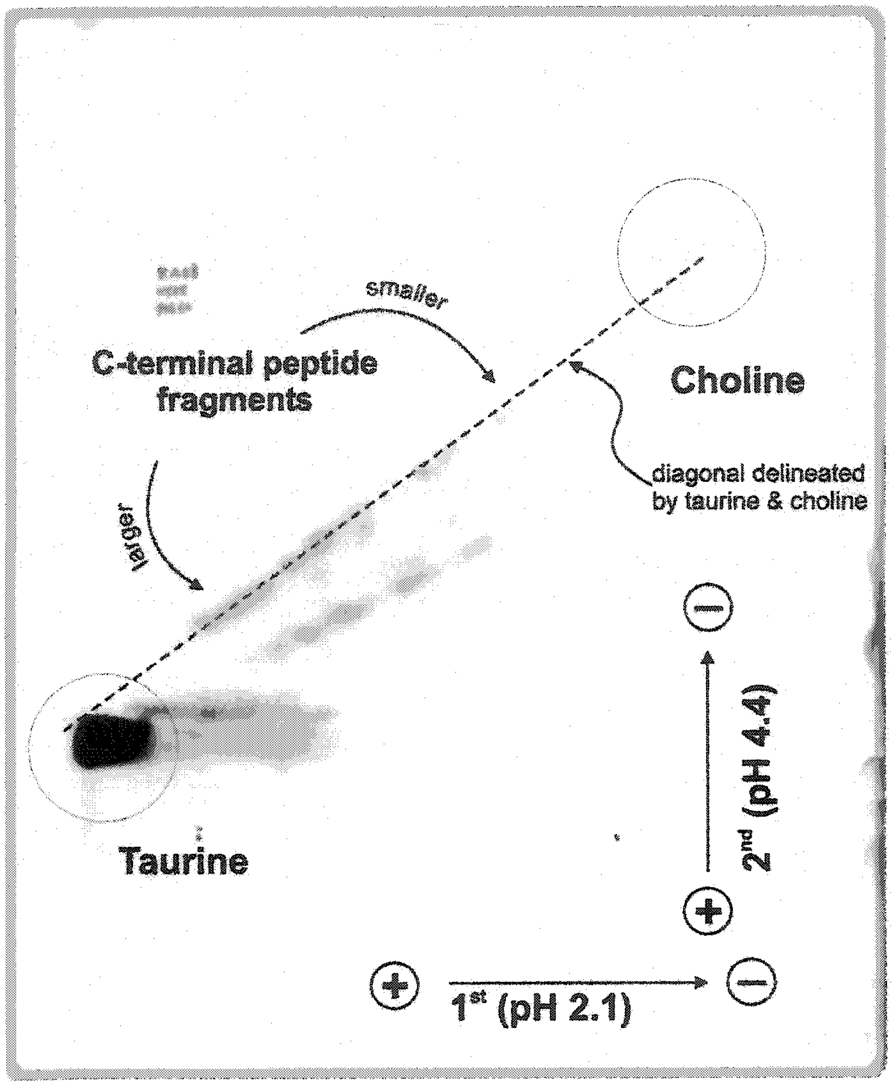


Figure 2-4 Typical 2D-HVPE diagonal.

Papain digested ^{14}C-methylated αCT (50:1 substrate: enzyme ratio) for 60 min. at 37°C. Radioactive markers taurine and choline are identified as spots at the origin and extremity of diagonal, respectively. The large C-terminal peptide fragments have low electrophoretic mobility, and are indicated with solid arrows. Dashed line delineates the diagonal containing the C-terminal peptides.

The dark spots observed on the X-ray film correspond to the radiolabelled peptides and markers (^{14}C-taurine and ^{14}C-choline) as described. For each digestion times a diagonal is clearly visible. Both enzymes are suitable for producing peptide fragments from the methylated protein. The papain digest seems to be the better enzyme in this case and a digest of 60 minutes seems sufficient in producing the most amount on-diagonal peptides based on their intensity. The peptides with lower electrophoretic mobility on the diagonal correspond to C-terminal peptide fragments of approximately 8 to 10 residues in length (with charge of +1). This length is more than enough to gain sequence information for application to database searching and other characterization requirements.

^{14}C trace methylation was used for visualisation purposes only. An autoradiogram of a papain digest of "cold" CH_3/CD_3 methylated αCT is shown in Figure 2-7. The diagonal, delineated by choline and taurine, was cut out as a strip of paper and the C-terminal peptides were eluted off with pH 2.1 buffer for ESI MS analysis. Unfortunately no peaks could be found corresponding to the double labelled peptides. This is most probably due to the impurities present in the Varsol used as insulator and coolant in the electrophoresis apparatus. The elution of the diagonal peptides from the paper was verified using the diagonal of the papain digested ^{14}C-αCT for 60 min. A first elution with Varsol was tried to remove any Varsol-soluble materials that were transferred onto the electrophoretic paper from the Varsol in the HVPE tank and the peptides were then eluted with pH 2.1 buffer. Radioactivity count of the eluted solutions confirmed that the elution step was efficient ($\text{Count}_{\text{Varsol}}$=33; $\text{Count}_{\text{pH 2.1}}$=7732) but an HPLC of the eluted peptides showed many highly absorbing peaks not consistent with peptides (data not shown) suggesting that Varsol is not an ideal solvent for eluting off these impurities.

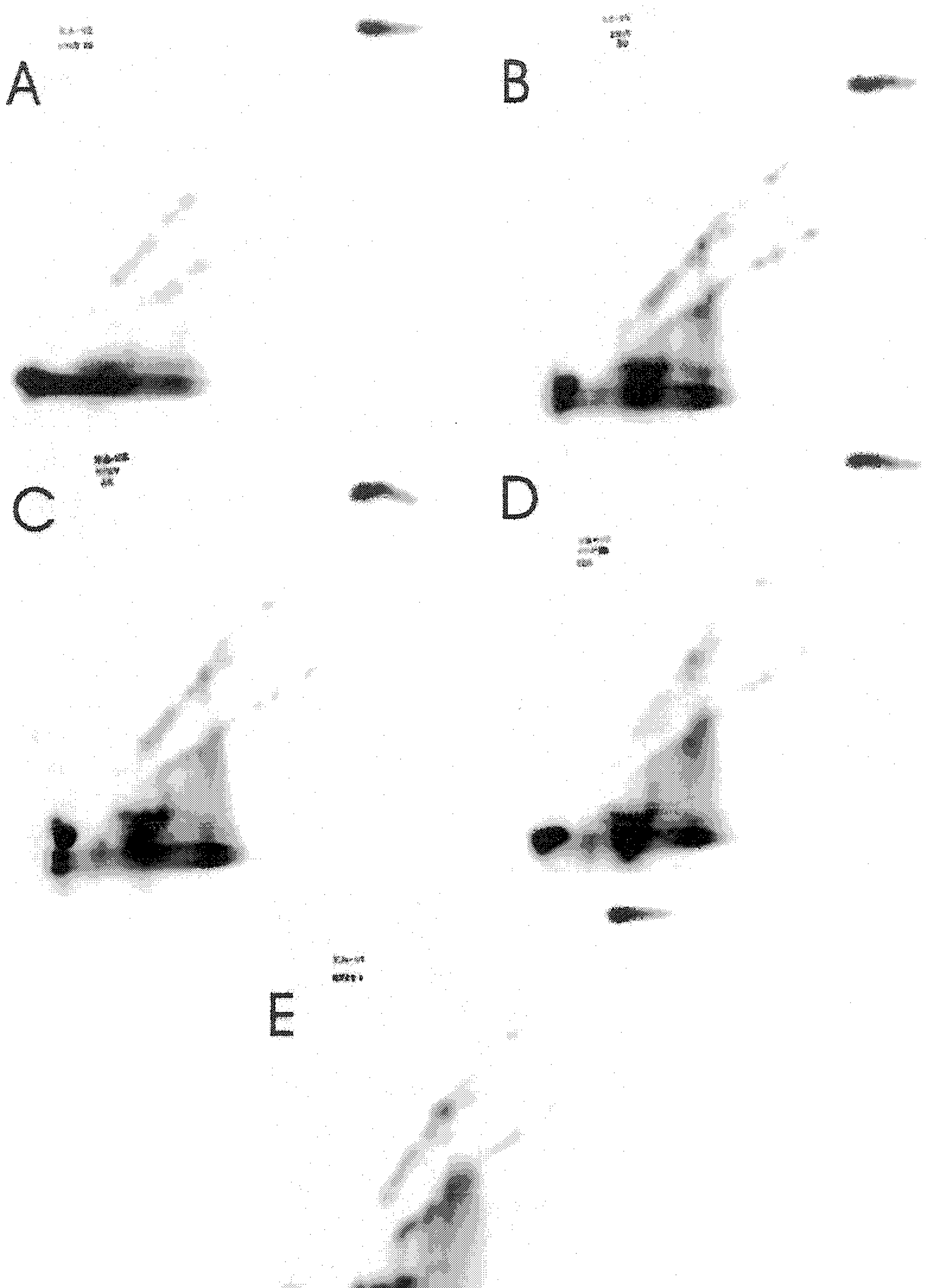


Figure 2-5 Autoradiograms of 2D-HVPE of a pepsin digest of ^{14}C-methylated αCT. Pepsin digest at 100:1 protein:enzyme ratio at 37°C. A) digestion for 15 min., B) 30 min., C) 60 min., D) 2 hr., and E) 24 hr.

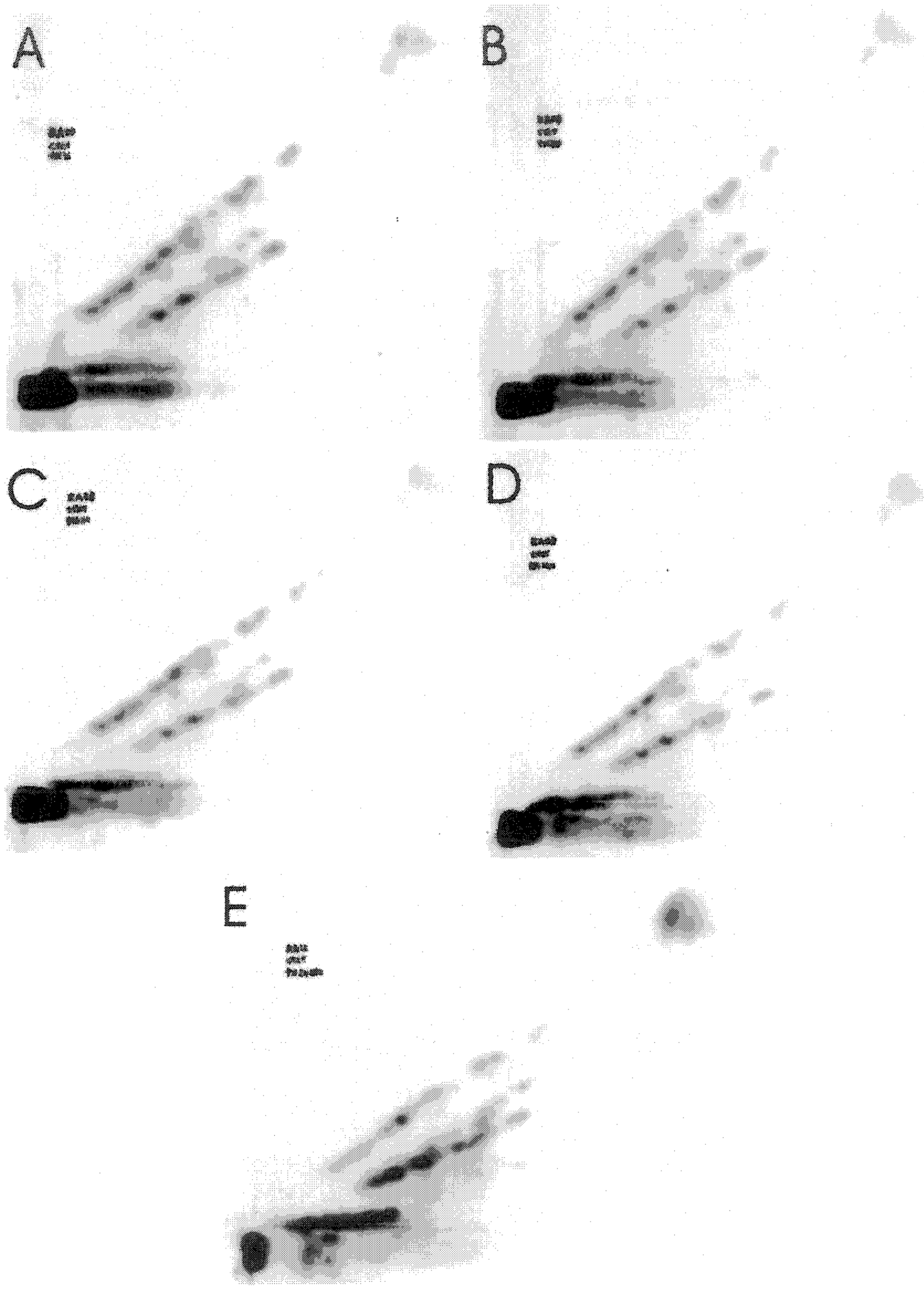


Figure 2-6 Autoradiograms of 2D-HVPE of a papain digest of ^{14}C-methylated αCT. Papain digest at 50:1 protein:enzyme ratio at 37°C. A) digestion for 15 min., B) 30 min., C) 60 min., D) 2 hr., and E) 24 hr.

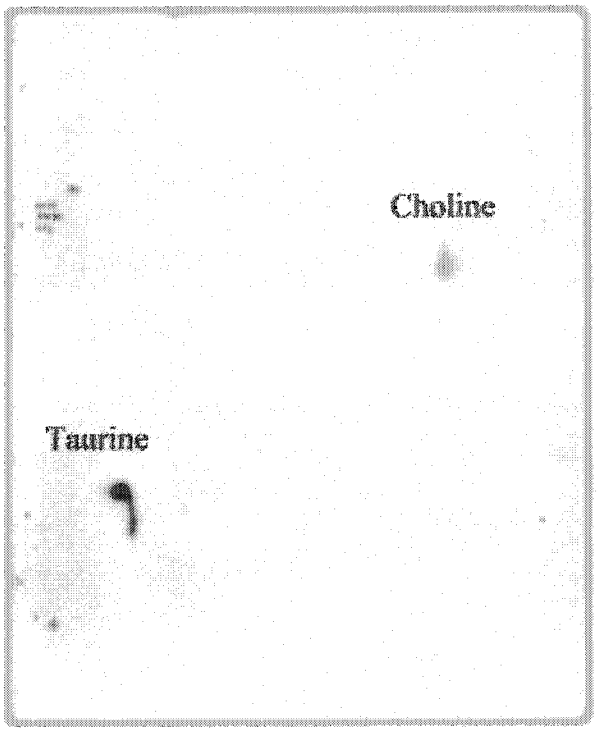


Figure 2-7 Autoradiogram of 2D-HVPE of CH_3/CD_3 methylated αCT. Protein was digested with papain for 30 minutes at 37°C (50:1 substrate:enzyme ratio). C-terminal peptides are located on the diagonal delineated by the taurine and choline markers.

The digestion of the αCT was carried out using two different proteases, namely, pepsin and papain. The use of pepsin has two advantages: the enzyme is active at low pH, so that prior denaturation of the substrate is rarely required, and also the disulphide interchange reactions are very slow at low pH (Allen, 1981). Papain is a thiol protease, and it is best reserved for the subfragmentation of peptides of up to 100 residues in length (Allen, 1981). Under the same digesting condition, papain seemed to be more effective than pepsin as indicated by the more intense and more defined spots on the autoradiograms. The enzymatic digestions were also carried out for different times of digestion with the objective of determining the optimal time interval for obtaining the largest C-terminal peptide fragments. At different durations of enzymatic digestion, the positions of the on-diagonal peptides and their quantities were variable, indicating multiple cleavage sites for the methylated protein by the enzyme.

Another notable advantage is that the current procedure can be carried out with relatively small amount of proteins, with the isolation of C-terminal peptides requiring about 0.5 mg of protein (*e.g.* 20 nmol for αCT). The autoradiograms revealed intense spots, indicating that the current procedure can in fact be carried out on significantly less material, but perhaps increasing the exposure time to the X-ray film for the detection of the labelled peptides in order to maintain the same degree of sensitivity.

A smaller apparatus was designed out of an existing horizontal gel electrophoresis system for agarose gel applications (see Figure 2-8) to overcome the problem with Varsol and to miniaturize the procedure for smaller amounts of digested protein. Whatman 3MM paper was used and cut to size (see Figure 2-9).

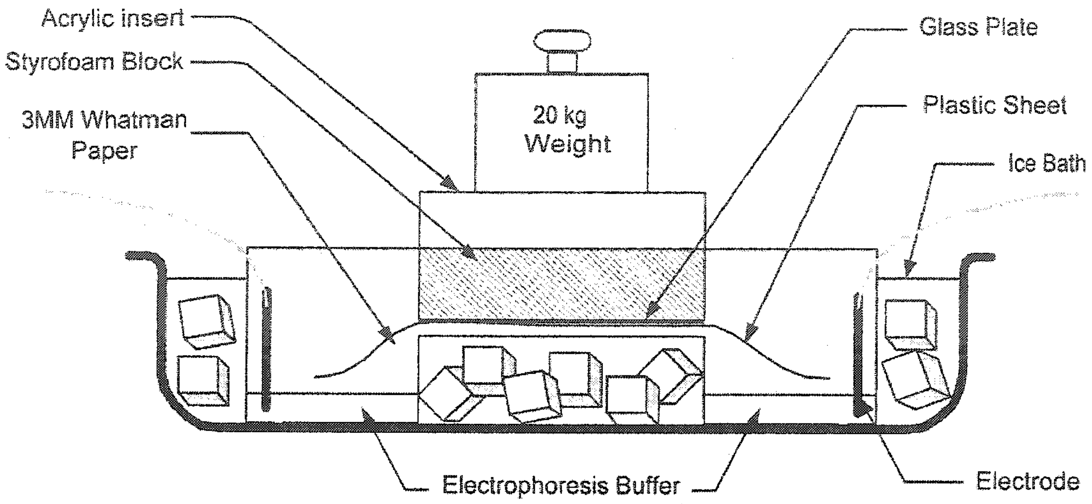


Figure 2-8 Diagram of small HVPE apparatus using an Owl EasyCast™ horizontal minigel electrophores1is system

To delineate the diagonal fluorescent markers were prepared to circumvent the use of radioisotopes. The use of fluorescent markers was chosen for rapid determination of a diagonal without the use of staining or radiolabelled markers. However, certain criteria must be fulfilled for the successful choice of fluorescent markers. Many highly fluorescent compounds are either salts or exhibit poor water solubility. Both of these properties are not compatible with HVPE. Markers must also possess the same electrophoretic mobility at both pHs employed to lie within the diagonal. Surprisingly many fluorescent compounds possess a pK_a near or below pH 4.4 (*e.g.* the pK_a of the dimethylamino group of Dns). It is for these reasons that the 7-methoxycoumarin-3-carboxylic acid moiety was chosen. Ethanolamine was chosen to give the neutral fluorescent marker as, although it is neutral, it demonstrated enough water solubility with its hydroxyl group. Spermidine was chosen arbitrarily for the simple reason that the coupling through one amine will result in a marker with a +2 net charge. The small

HVPE system developed was subsequently found to be similar in design to a commercially available apparatus sold under the name of the Hunter Thin Layer Electrophoresis (HTLE) apparatus which uses polyester backed cellulose sheets instead of the 3MM Whatman paper. This apparatus has been used for the separation of phospho-peptides and phospho-amino acids (Boyle *et al.*, 1991).

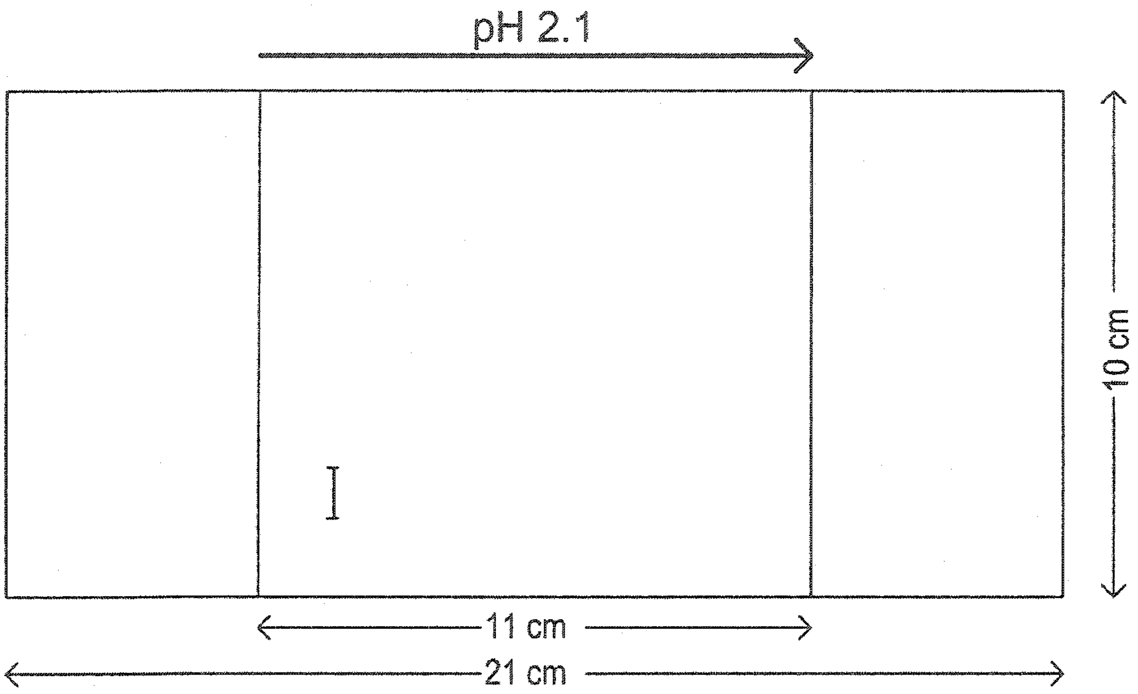


Figure 2-9 Dimensions of 3MM Whatman paper used in small HVPE apparatus

The apparatus performed well for a one dimension run and a typical result is shown in Figure 2-10. A carboxypeptidase/pronase digest of the peptide H-SASAFA-OH (as described in materials and methods) was run with Ser, Ala and Phe standards and flanked by Dns-Arg. A 2D run was also successful on the tripeptide methyl ester H-AAA-OCH_3 (see Figure 2-11) with the diagonal delineated by the coumarin markers (visible under UV) and on the pepsin digest of *in vacuo* ^{14}C trace methylated αCT with

^{14}C-MeOH-HCℓ (see Figure 2-12). The radiolabelled diagonal markers, choline and taurine are clearly visible with peptides found on the diagonal.

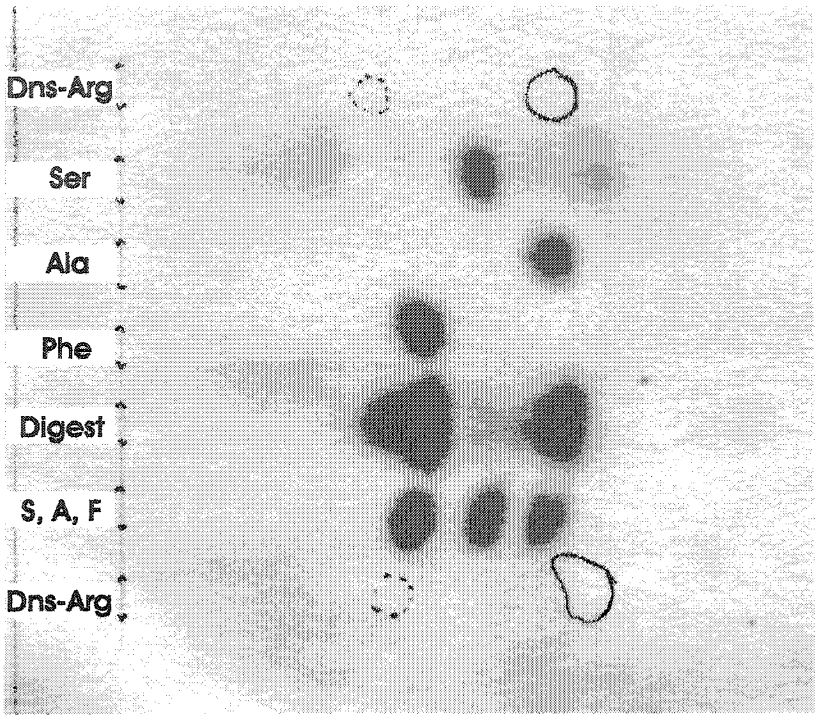


Figure 2-10 Ninhydrin stained small HVPE of carboxypeptidase digest of H-SASAFA-OH with Ser, Ala and Phe standards and flanked by Dns-Arg

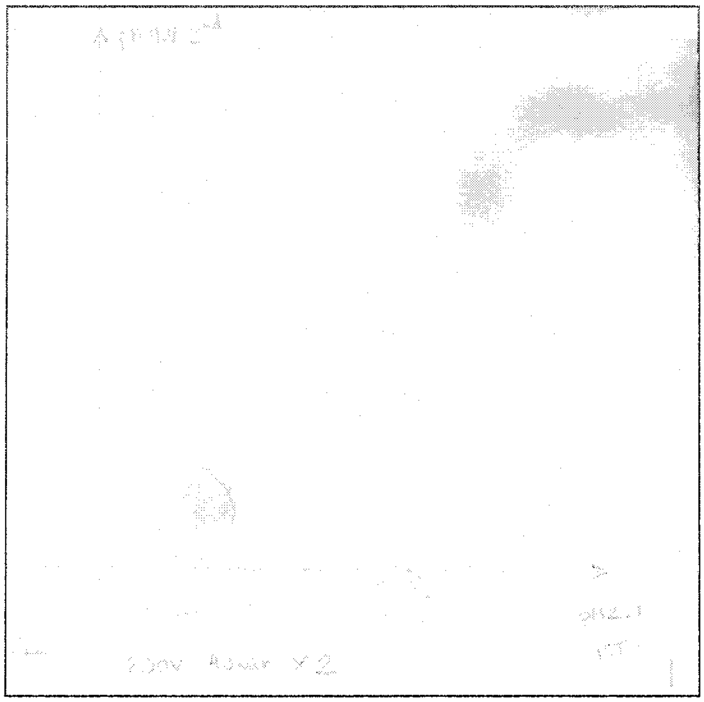


Figure 2-11 Ninhydrin stained small 2D-HVPE diagonal of H-AAA-OCH_3. Diagonal delineated by coumarin markers.

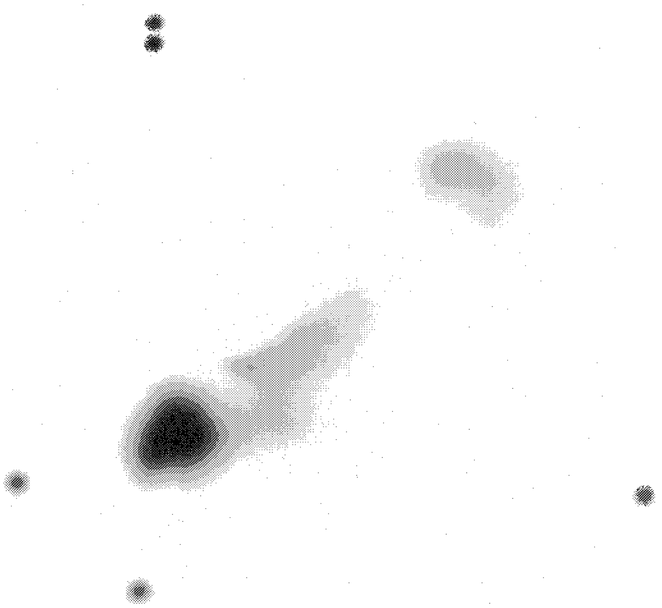


Figure 2-12 Small 2D-HVPE autoradiogram of a pepsin digest of αCT methylated *in vacuo* with ^{14}C-MeOH-HCℓ. Diagonal delineated by ^{14}C-choline and ^{14}C-taurine markers.

In summary, a proof of concept for the use of the *in vacuo* methylation procedure to label the C-terminal end(s) of a protein for subsequent separation and isolation by a "diagonal" electrophoretic method has been shown. This method demonstrates a significant improvement over previous methods as, although relatively large amounts of protein (mg) were used for visualization purposes, such large amounts are not necessary with this method, and lends itself well to the analysis of small amounts of protein (sub-μg). The *in vacuo* methodology surpasses the pre-existing aqueous approach as only a small amount of reagent is needed, excess reagent can be re-used and is simply removed by breaking the vacuum and further purification of the methylated protein is not required. A high amount of radiolabelling was accomplished employing this method with a small amount of reagent displaying the efficiency of the *in vacuo* chemical modification approach. A miniature equivalent HVPE apparatus was developed and successfully used to accomplish the separation of C-terminal peptides. This smaller apparatus provides a faster and less labour intensive method as compared to the previous larger vertical apparatus. With the use of fluorescent markers to delineate the diagonal, C-terminal peptides can be eluted off the paper "blindly" for MS/MS analysis to gain sequence information. Although final analysis with MS of the peptides from the large electrophoresis system was unsuccessful and no attempt was made for MS analysis of peptides separated on the miniature system, a proof of concept was nevertheless accomplished. The use of paper may also be avoided by using the polyester backed cellulose sheets commonly used with the HTLE apparatus. Where the sequential degradation methods have failed in the case of a blocked C-terminal, this approach has shown to work effectively in the case of a protein either with unblocked C-terminal (native chymotrypsin) or with blocked C-terminal (the methyl-ester derivative of chymotrypsin). The approach taken therefore has a wider range than existing methodologies in terms of applicability.

Chapter 3 *In vacuo* methylation of imprinted lyophilized α-chymotrypsin

Introduction

The fact that lyophilized enzymes have substantial catalytic activity in anhydrous organic solvents (Zaks and Klibanov, 1988; Chen and Sih, 1989; Klibanov, 1989; Wescott and Klibanov, 1994; Koskinen and Klibanov, 1996) implies that a portion of the active-sites retain an active conformation under nonaqueous conditions. Active-site titration data (Zaks and Klibanov, 1988; Wangikar *et al.*, 1995) and solid-state NMR (Burke *et al.*, 1992) studies of lyophilized chymotrypsin indicate that the active-site of the enzyme consists of a mixture of active and inactive conformations. The addition of ligands to the aqueous solution prior to lyophilization has been shown to increase the catalytic activity of the lyophilized enzyme in organic solvents (Russell and Klibanov, 1988; Dabulis and Klibanov, 1993; Triantafyllou *et al.*, 1997). In the case of chymotrypsin, lyophilization in the presence of competitive inhibitors had a substantial activation effect (Zaks and Klibanov, 1988). Other multifunctional ligands such as sorbitol which are structurally unrelated to substrates also have an activating effect (Dabulis and Klibanov, 1993; Triantafyllou *et al.*, 1997) and appear to act as lyoprotectants maintaining the conformational integrity of the active-site. It has been postulated (Russell and Klibanov, 1988; Yennawar *et al.*, 1995; Griebenow and Klibanov, 1997) that the binding or association of such ligands imprints (Braco *et al.*, 1990; Staahl *et al.*, 1991; Klibanov, 1995; Yennawar *et al.*, 1995; Mishra *et al.*, 1996) a catalytically favourable conformation in the catalytic site of lyophilized enzymes. This hypothesis implies that the chemical properties of active-site functional groups are altered in the imprinted enzyme. The following describes an approach in using the *in vacuo* chemical modification of lyophilized proteins (Kaplan and Taralp, 1997; Taralp and Kaplan, 1997) to investigate the chemical reactivity to iodomethane of the active-site imidazole of His-57 of α-chymotrypsin (αCT) lyophilized in the presence and absence of imprinting ligands.

Materials and Methods

Bovine αCT (EC 3.4.21.1) (lot#91H7195), iodomethane, *N*-acetyl-L-tryptophan, *N*-acetyl-D-tryptophan, indole, D-sorbitol and *N*-acetyl-L-tyrosine ethyl ester (ATEE) were purchased from Sigma Chemical Co. Sodium citrate was obtained from Fisher Scientific Co. All other reagents and chemicals were high purity preparations obtained from commercial sources.

Sample preparation

Stock solutions (20 mM) of ligands were adjusted to pH 8.0 with dilute NaOH. An aliquot (5.0 mL) of a ligand stock solution was added to 1.0 mL of α-chymotrypsin (1.0 mg/mL) in dH_2O and the volume adjusted to 10.0 mL with dH_2O to give a solution containing αCT (100 μg/mL) and ligand (10.0 mM). Each αCT-ligand solution was adjusted to pH 8.0 with dilute NaOH, transferred in 1.0 mL aliquots to Pyrex® glass hydrolysis tubes and lyophilized.

In vacuo reaction with iodomethane (25 μL) was carried out at 75°C as previously described (Kaplan and Taralp, 1997; Taralp and Kaplan, 1997) with some minor modifications to the method. The use of a side arm was not necessary as the iodomethane would freeze to the sides of the reaction vessel as it was introduced preventing it from "wetting" the lyophilized protein. A general representation of the method is described in Figure 0-1. A nitrogen line was not used, although this was added to the procedure in later experiments to minimize the introduction and condensation of water from the air. At various time points the reaction was stopped by trapping out the iodomethane and breaking the vacuum by applying the tip of a white hot glass rod to the scored reaction vessel. The samples were stored *in vacuo* over P_2O_5 until they were assayed.

Sample preparation of active site protected αCT for NMR analysis

αCT (20 mg) was dissolved with or without *N*-acetyl-L-tryptophan (10 mM), the pH adjusted to 8.0 with the addition of 1M NaOH via micrometer-syringe and the final volume adjusted to 200 mL. Samples were lyophilized in round bottom flasks (500 mL), the lyophilized protein samples were transferred to large borosilicate hydrolysis tubes (25 x 200mm) and methylated *in vacuo* with iodomethane for 48 hours at 75°C as previously described. The methylated protein samples were re-dissolved in ddH_2O (ca. 20 mL) and dialyzed (MWCO 3500) against ddH_2O (4 x 4L). Samples were lyophilized, transferred to borosilicate hydrolysis tubes and methylated *in vacuo* with ^{13}C iodomethane.

Protein samples were dissolved in a total of 1 mL (8M urea 80% D_2O, 200 mM NaH_3PO_4 pH 7.5) with the addition of acetonitrile (30 μL) as internal standard. Spectra were recorded on a Bruker XL 300 with a total of 6400 transitions.

Quantification of enzymatic activity

Each sample was dissolved in pH 4.0 sodium formate (100 mM, 1.0 mL) buffer prior to assay. Enzymatic activity was quantified using a pH stat (Titration 11/Ole Dich autotitrator assembly) set to pH 8.0 on aliquots of samples (100 μL) by following the rate of hydrolysis of ATEE (10.0 mM ATEE, 0.10 N KCℓ and 5% (v/v) acetonitrile, 5mL). Initial velocities were recorded as the reaction was stirred with a magnetic stir bar at 20°C under a nitrogen atmosphere with 0.100 M NaOH as titrant introduced *via* micrometer glass Agla™ syringe.

Results and Discussion

αCT lyophilized at LpH 8.0 in the presence or absence of ligands retained full activity after heating *in vacuo* for 12 h at 75°C (see Table 3-1). Under the same conditions, the ligand-free lyophilized enzyme lost most of its catalytic activity after *in vacuo* methylation for 12h with iodomethane (33% activity remaining).

Although lyophilization in the presence of ligands did not alter the thermal stability of the enzyme, it did have significant effects on the inactivation by iodomethane. The L and D isomers of *N*-acetyltryptophan effectively protected the lyophilized enzyme from inactivation. Indole did not protect, but surprisingly accelerated the rate of inactivation. Citrate and sorbitol which have no structural similarity to any substrate of chymotrypsin had strikingly distinct but opposite effects. Citrate effectively prevented inactivation whereas sorbitol dramatically accelerated the inactivation.

Table 3-1 Catalytic activity of α-chymotrypsin after *in vacuo* reaction with iodomethane of the enzyme co-lyophilized with or without ligand/excipient[a]

Ligand/ Excipient 10 mM	Heated Control Time (h)					Reaction with Iodomethane Time (h)			
	0	2	4	8	12	2	4	8	12
None	1.0	1.0	0.99	1.1	0.99	0.63	0.54	0.41	0.33
N-Ac-L-Trp	1.1	1.1	1.2	1.2	1.0	1.1	1.1	0.98	0.99
N-Ac-D-Trp	1.1	1.1	1.0	0.97	0.93	1.0	1.1	1.0	0.95
Indole	0.90	0.91	1.0	0.97	0.93	0.36	0.33	0.22	0.08
Citrate	1.0	0.99	0.97	1.0	1.0	0.98	1.0	1.0	0.99
Sorbitol	0.99	0.94	0.98	0.87	0.90	0.13	0	0	0

[a] The activities are relative to unheated αCT lyophilized at pH 8.0. The values reported are the average of two independent experiments. Mean deviations in all cases were less than 10%.

The histidine residues of lyophilized α-chymotrypsin are dimethylated by reaction with iodomethane octane or *in vacuo* at LpH 8.0 (Taralp and Kaplan, 1997). Chemical analysis and ^{13}C-NMR spectra after reaction of proteins with ^{13}C-iodomethane for 24 h at 75°C shows complete reaction with imidazole side-chains and less than 20% reaction with amino groups (Taralp and Kaplan, 1997). Acetylation of the ϵ-amino groups does not affect the enzymatic activity of chymotrypsin (Oppenheimer *et al.*, 1966) and, similarly, reductive methylation, which retains the charge on the amino groups, has little or no detectable effect on the catalytic properties of bovine trypsin (Rice *et al.*, 1977). Similarly, methylation of amino groups to their trimethylammonium derivatives is expected not to significantly affect enzymatic activity. The observation that full activity is retained after methylation in the presence of citrate or *N*-acetyltryptophan is consistent with modification of amino groups not affecting the catalytic activity. It therefore appears that the inactivation of the enzyme is entirely due to dimethylation of the imidazole side-chain of the active-site histidine. Further evidence for this was obtained by differential labelling in which chymotrypsin co-lyophilized with *N*-acetyl-L-tryptophan (10 mM) was reacted with ^{12}C-iodomethane at LpH 8.0 for 48 h, followed by dialysis in water to remove the ligand. Subsequent reaction of the ligand-free ^{12}C-methylated lyophilized protein with ^{13}C-iodomethane gave the characteristic ^{13}C-resonances of dimethylhistidine (Taralp and Kaplan, 1997) at 34.08 and 36.59 ppm (see Figure 3-1) indicating that the imidazole of His-57 was protected by the bound ligand.

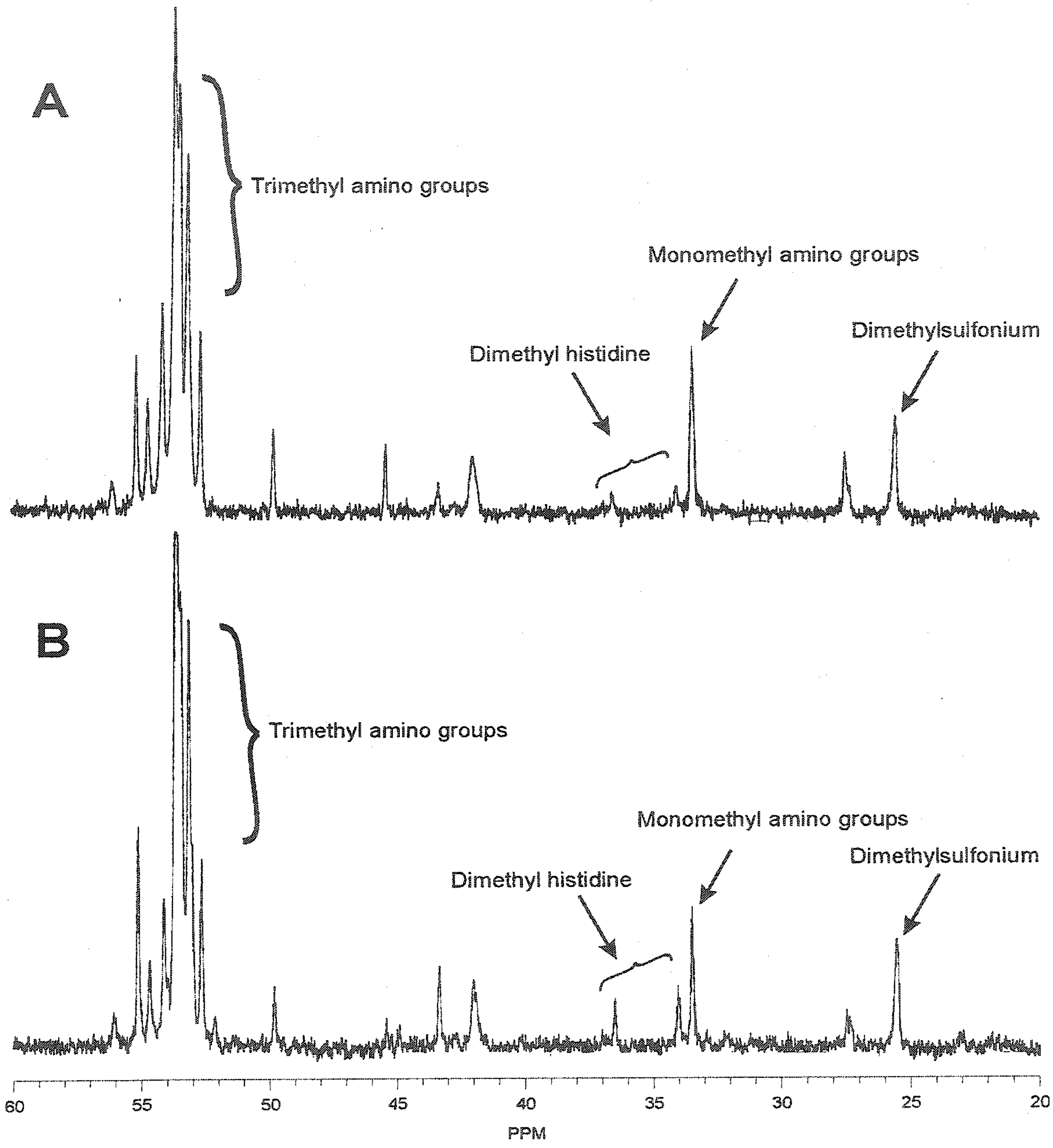


Figure 3-1 ^{13}C NMR comparison of *in vacuo* methylated αCT co-lyophilized without (**A**), or with (**B**) *N*-acetyl-L-tryptophan. Note the difference in intensity of the dimethyl histidine peaks.

αCT was co-lyophilized with (**B**) *N*-acetyl-L-tryptophan or (**A**) without a ligand and methylated *in vacuo* with ^{12}C iodomethane. The ligand was removed by dialysis and the protein was lyophilized and re-methylated *in vacuo* with ^{13}C iodomethane.

The remarkable increased rate of inactivation of lyophilized αCT with bound indole was unexpected. This must be due to the increased rate of methylation of a catalytically essential functional group. NMR evidence from previous studies (Kaplan and Taralp, 1997; Taralp and Kaplan, 1997) shows that the only group in the active-site that is methylated *in vacuo* by iodomethane is the imidazole of His-57. From the data in Table 3-1, it appears that the rate of inactivation of the ligand-free lyophilized chymotrypsin is faster in the early stages of the reaction. This observation can be explained by the existence of two populations of the enzyme active-site in the lyophilized enzyme (Figure 3-2) in which iodomethane reacts more rapidly with one population.

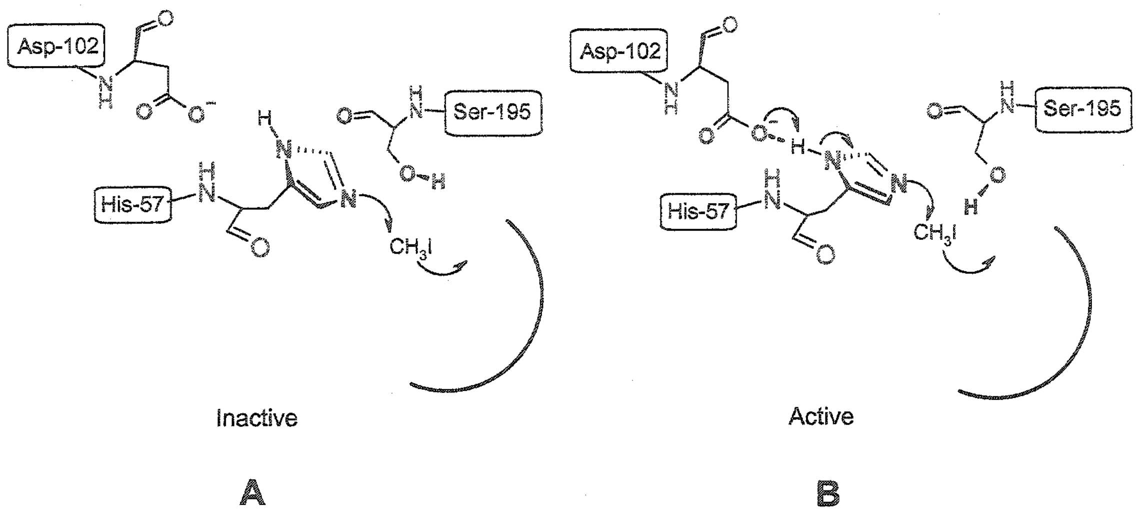


Figure 3-2 Schematic representation of the active site of lyophilized αCT in which two conformational populations exist. In **A** the imidazole group of His-57 does not interact with the Asp-102 carboxylate group and is less reactive than in **B** where the imidazole is hydrogen bonded to the carboxylate group

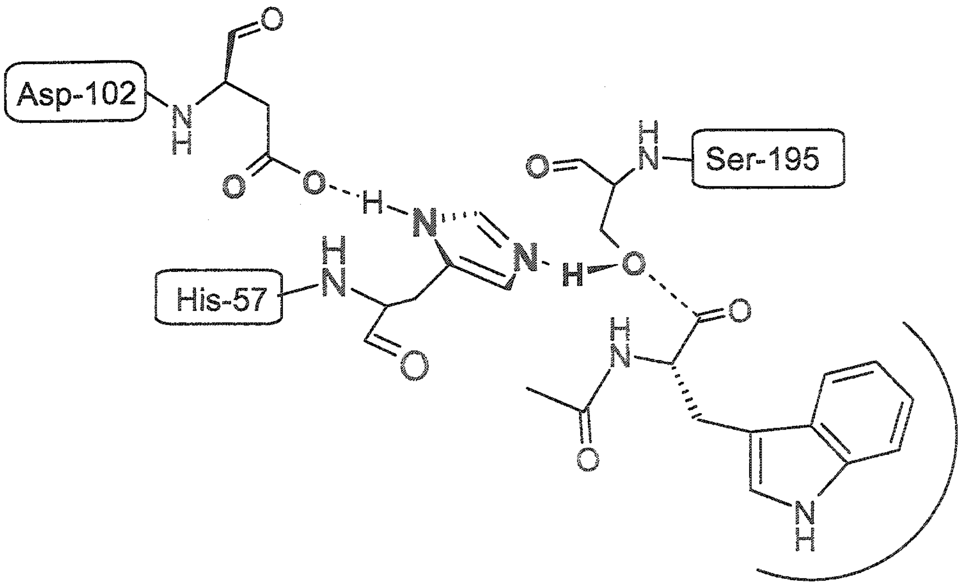


Figure 3-3 Schematic representation of the effect of binding of *N*-acetyl-L-tryptophan on the reactivity of the imidazole side-chain of His-57 in the active-site of lyophilized αCT

The ligand-free lyophilized enzyme consists of a mixture of the conformations A and B, but when indole is bound to the lyophilized enzyme a greater proportion of the active-site groups are present in conformation B. The imidazole group is made more nucleophilic by virtue of its interaction with the carboxyl group giving rise to a faster rate of reaction with iodomethane and this is reflected in a faster rate of inactivation of the enzyme.

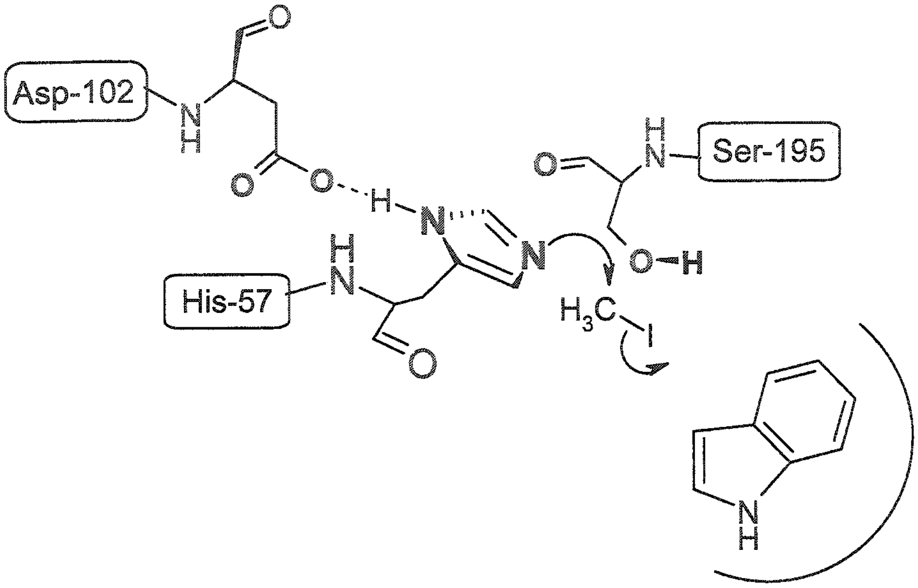


Figure 3-4 Schematic representation of the effect of binding of indole on the reactivity of the imidazole side-chain of His-57 in the active-site of lyophilized αCT

After 2h, the data fits a first order inactivation process (Figure 3-5), but clearly in the initial stages the reaction is occurring by a more rapid process. Extrapolation of the linear region to zero time gives an intersection of –0.35 on the y-axis, which corresponds to 30% of the active sites in the lyophilized enzyme being present in the most reactive form. These observations are consistent with active-site titrations (Zaks and Klibanov, 1988; Wangikar *et al.*, 1995) which indicate that the active-sites of lyophilized enzymes are a mixture of active and inactive conformations. In Figure 3-3, *N*-acetyl-L-tryptophan binds and shields the imidazole of His-57 preventing inactivation.

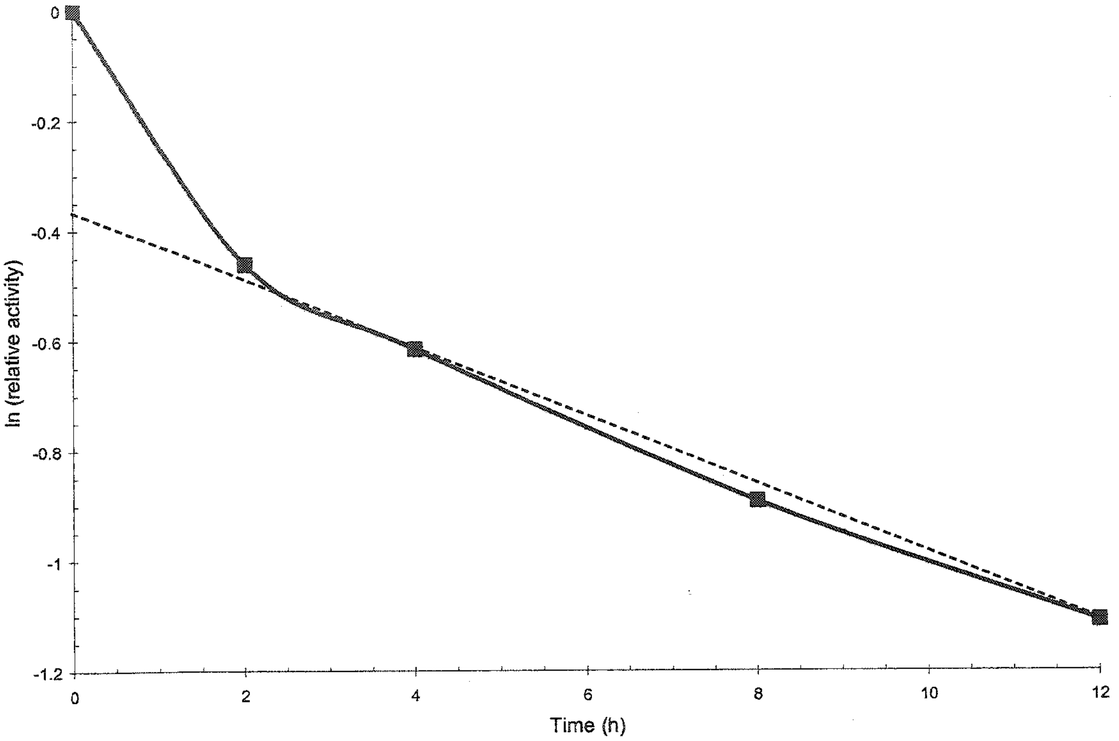


Figure 3-5 First order plot for the inactivation of lyophilized αCT by *in vacuo* methylation with iodomethane

The crystal structure of αCT indicates the active-site functional groups of Asp-102, His-57 and Ser-195 are in position to form a hydrogen-bonded catalytic triad (Tsukada and Blow, 1985). However, the crystal structure of γ-chymotrypsin indicates that this hydrogen-bonded system is altered at moderately alkaline pH values (Cohen *et al.*, 1981) such as the pH used in this experiment. Nevertheless, it is generally agreed that hydrogen bonding of the imidazole side-chain of His-57 with the carboxylate of Asp-102 and the hydroxyl of Ser-195 is a crucial component of the catalytic process (Blow *et al.*, 1969; Fersht, 1985; Creighton, 1993). The fact that the imidazole of His-57 reacts so readily with iodomethane in the absence of bound substrate provides further evidence that the side chains of Ser-195 and His-57 are not strongly hydrogen bonded (Steitz and Shulman, 1982). Therefore, one can propose that the catalytic triad of αCT at LpH 8.0 is found as two or more hydrogen-bonded forms of the imidazole of His-57 with the carboxylate of Asp-102, resulting in two different populations of lyophilized enzyme (see Figure 3-2). Due to the conformational rigidity of the lyophilized protein (Rupley and Careri, 1991; Broos *et al.*, 1995; Xu and Klibanov, 1996; Klibanov, 1997), conformations such as that depicted in Figure 3-2A where the imidazole of His-57 is not positioned to interact optimally with carboxylate Asp-102 would be catalytically less active or inactive. When lyophilized in the presence of a ligand which binds in the active-site but not with the reactive functional groups of the active-site (*e.g.* indole), a greater proportion of the active-sites in the lyophilized enzyme are maintained in a catalytically active conformation (Figure 3-2B and Figure 3-4) where the imidazole moiety is optimally hydrogen-bonded to the carboxylate of Asp-102. Such an interaction would be expected to increase the nucleophilicity of the imidazole group and therefore increase the rate of reaction with electrophilic reagents such as iodomethane. In fact, the active-site imidazole of His-57 of aqueous chymotrypsin (Cruickshank and Kaplan, 1972) and chymotrypsinogen have been found to be super-reactive toward the electrophilic reagent 1-fluoro-2,4-dinitrobenzene (Cruickshank and Kaplan, 1974).

These results provide further evidence for the existence of an imprinting phenomenon which affects the structure of the active-site and are consistent with hydrogen bonding between His-57 and Asp-102 as part of this process. It is possible that *N*-acetyltryptophan also induces the formation of a hydrogen bond of the active-site imidazole with the hydroxyl of Ser-195, as has been proposed for substrates (Steitz and Shulman, 1982), thereby preventing reaction with iodomethane (see Figure 3-3). The other possibility is that the amino acid functionality simply shields the catalytic groups and prevents reaction with iodomethane. Whichever explanation is correct, it is clear that *N*-acetyltryptophan imprints the lyophilized enzyme by affecting the hydrogen bonding of the catalytic triad.

The reason citrate prevented and sorbitol accelerated the inactivation is not obvious. Both are multifunctional molecules that act as excipients and are believed to interact, primarily through hydrogen bonding, with the protein during the lyophilization process to minimize protein-protein contacts (Costantino *et al.*, 1995; Mishra *et al.*, 1996; Griebenow and Klibanov, 1997). The present data indicates that the association of citrate with the lyophilized enzyme shields the catalytic site and prevents methylation of His-57. In contrast, it appears that sorbitol does not shield the active-site but does interact with other regions of the molecule in such a manner as to maintain the active-site of the lyophilized enzyme in a catalytically active conformation. Like indole, it does not prevent methylation of His-57, but it imprints α-chymotrypsin with a higher proportion of the active-sites in an active conformation (Figure 3-2B), and this is reflected in a more rapid inactivation by iodomethane. Sorbitol has many interaction sites with the protein that provide a much greater stabilization of the native structure and the active-site in the lyophilized state.

A significant difference between lyophilized protein and protein in solution is that there are no dynamic conformational or protonic equilibria in the lyophilized state. This gives rise to two phenomena. One is "pH memory" (Zaks and Klibanov, 1988; Klibanov, 1995; Xu and Klibanov, 1996; Klibanov, 1997) (*vide supra*) and the other is molecular memory (Braco *et al.*, 1990; Staahl *et al.*, 1991; Klibanov, 1995; Yennawar *et al.*, 1995; Mishra *et al.*, 1996) in which the protein "remembers" the conformation imprinted by a ligand co-lyophilized with the protein. In the present study, we have taken advantage of these phenomena to carry out a chemical reaction on chymotrypsin that could not be performed in aqueous solution. Preparing the lyophilized protein at LpH 8.0 greatly increased the selectivity towards modification of the imidazole groups since the majority of the amino groups were protonated and therefore unreactive towards iodomethane. In aqueous media, the L and D enantiomers of *N*-acetyltryptophan are relatively weakly bound to chymotrypsin with association constants of the order of 10^2 (Barman, 1969). In solution, they would offer little protection to chemical modification as a significant portion of the free enzyme would be present in equilibrium with the enzyme-inhibitor complex and, due to Le Chatelier's principle, the active-sites would become completely modified.

The effect of indole and sorbitol in increasing the apparent rate of methylation of the active-site histidine indicates that the imprinting phenomena are involved in orienting the catalytic triad of chymotrypsin for optimal catalytic activity. Lyophilization in the presence of some inorganic salts has also been reported to dramatically increase the activity of enzymes in organic solvents (Khmelnitsky *et al.*, 1994), and presumably, they affect the conformation of the active-site when co-lyophilized with the enzyme. It was also found in preliminary investigations that the presence of various inorganic anions and cations alters the rate of reaction of iodomethane with functional groups in lyophilized αCT and other proteins, however their effects appear to be more complex than with organic ligands. In the presence of 5 mM phosphate, for example, *N*-acetyl-L-tryptophan still protects αCT from inactivation (Taralp and Kaplan, 1997) by reaction with iodomethane but the protection is not 100% as in the present investigation where no anions or cations were added before lyophilization.

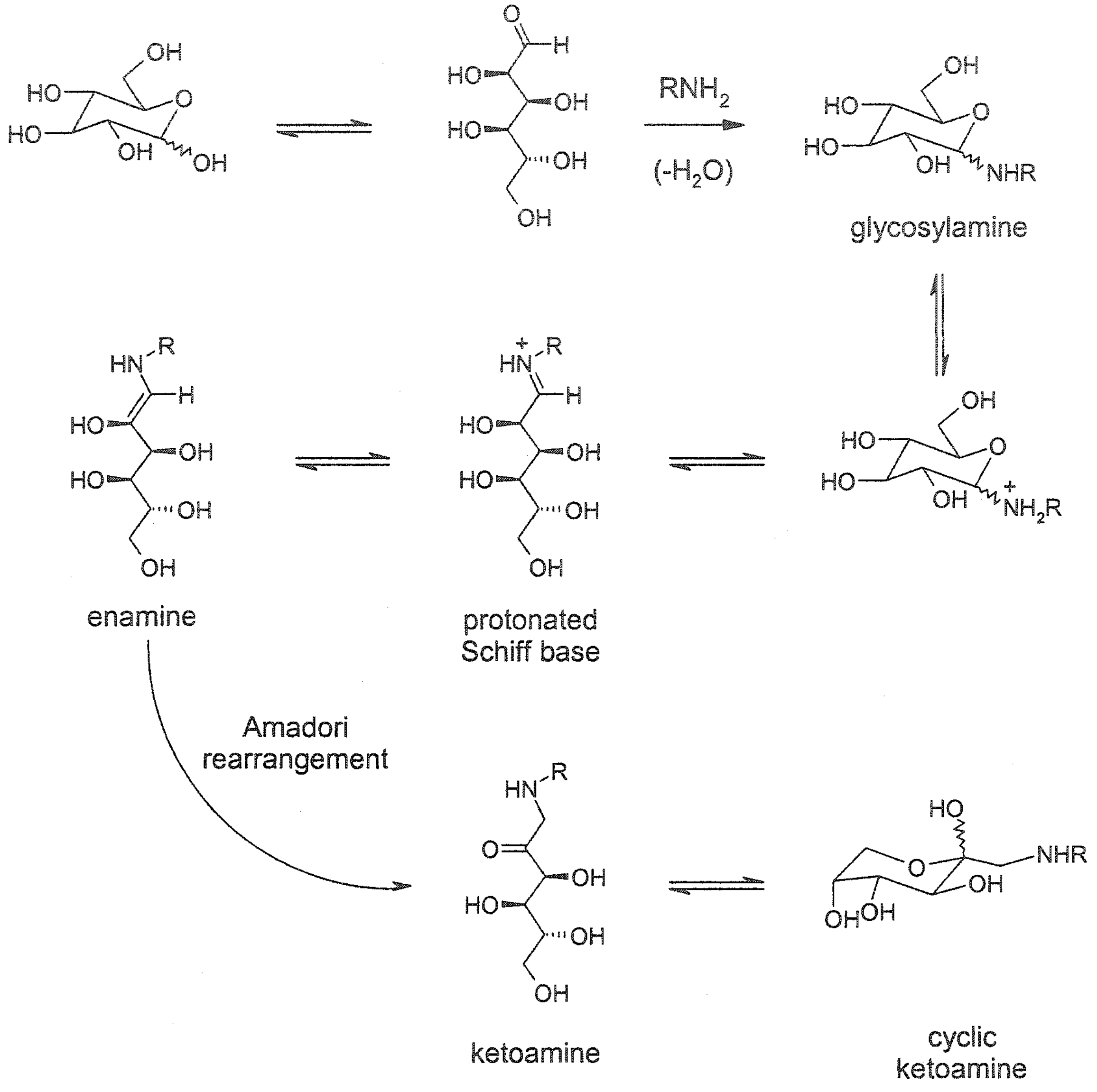
In summary, the results provide evidence, through chemical modification, that the imprinting phenomenon (Braco *et al.*, 1990; Staahl *et al.*, 1991; Klibanov, 1995; Yennawar *et al.*, 1995; Mishra *et al.*, 1996) on lyophilized αCT affects the chemical properties of the active-site by inducing a catalytically favourable orientation and interaction of the active-site functional groups during lyophilization. The results we have obtained so far suggest that ligands, which interact to stabilize the aqueous protein structure upon lyophilization, either enhance or abolish the chemical reactivity of active-site groups. Whether or not these observations turn out to be true generally, nonaqueous chemical modification of proteins lyophilized in the presence of ligands should provide the opportunity to determine the chemical properties of enzyme active-sites and to carry out specific chemical modifications that could not be achieved in aqueous solution.

Chapter 4 *In vacuo* glycation of lyophilized proteins

Introduction

Glycosylation of proteins plays an important role in the study of the theoretical and practical aspects of protein structure-function relationships (Opdenakker *et al.*, 1993; Rudd and Dwek, 1997; Bertozzi and Kiessling, 2001). The term "glycosylation", though sometimes loosely applied to processes involving the attachment of carbohydrates to other species, is more correctly confined to those cases where the linkage established is in fact glycosidic. Glycoproteins are involved in a variety of biological functions including antigen binding to antibody molecules, ligand binding by cellular adhesion molecules, cryoprotection, enzymatic activity and immunogens; as such, can be used as diagnostic reagents, and artificial vaccines. Most glycosylation in mammals occurs according to the familiar route of glycosyltransferases operating on membrane-associated substrates in the endoplasmic reticulum and Golgi apparatus and this type of enzyme-mediated glycosylation is substrate specific leading to selective glycosylation. However, other means of adding carbohydrates to proteins exist, which are not mediated by enzymes. This nonenzymatic "glycosylation", called glycation, involves the non-selective chemical attachment of reducing sugars to lysine residues or N-termini of peptides. Over 90 years ago, Louis Maillard reported that amino acids heated in the presence of reducing sugars resulted in the formation of brown pigments: "melanoidins". In this reaction an amine reacts with a reducing sugar to form a glycosyl amine. This amine then rearranges via the Amadori rearrangement to form 1-amino-2-keto sugar which tautomerizes and forms coloured bodies (Maillard, 1912). These coloured bodies impart flavours and aromas to food during cooking. Unfortunately, from a synthetic point of view, this reaction is not favourable as it initiates a complex biochemical cascade commonly known as the Maillard reaction.

This reaction (see Scheme 4-1) involves, initially, the nucleophilic attack of the protein's α or ε free amino groups on the anomeric carbon of the open chain form of glucose. A subsequent loss of water results in the formation of an aldimine derivative (Schiff base) for which the equilibrium is unfavourable but the Schiff base slowly undergoes Amadori rearrangement (Yaylayan and Huyghues-Despointes, 1994) to yield a relatively stable derivative, commonly termed the keto-amine adduct (Neglia *et al.*, 1983; Neglia *et al.*, 1985; Watkins *et al.*, 1985). Glycation in aqueous media proceeds *via* the formation of a Schiff base between the reducing sugar and a suitable α or ε-amino group (see Scheme 4-1), which encompasses all possible reactions that generate 'advanced glycation end-products' (AGEs) (Monnier *et al.*, 1996) from the comparatively unstable Schiff base. The accumulation of AGEs, due to age, in collagen and other long-lived proteins has been implicated in a range of degenerative diseases including arteriosclerosis (Vlassara, 1996), Alzheimer's (Colaco *et al.*, 1996; Thome *et al.*, 1996; Munch *et al.*, 1997) and the progressive complications of diabetes mellitus (Nakamura *et al.*, 1993; Bucala *et al.*, 1994; Paul and Bailey, 1996). For example, glycation of collagen due to diabetic hyperglycaemia affects the interaction between collagen and other proteins and this in turn affects the flexibility and permeability of tissues. Additionally, structural changes in glycated hemoglobin may contribute to hindered oxygen release under tissue hypoxia (Watala *et al.*, 1996).



Scheme 4-1 Reaction scheme of solution glycation; the first step in the Maillard reaction

Glycation has been observed to occur with lyophilized peptides (Tarelli *et al.*, 1994) and proteins (Li *et al.*, 1996; Quan *et al.*, 1999; Yeboah *et al.*, 1999) when a reducing sugar is in the formulation as an excipient for the stabilization of the dried product. The synthetic addition of carbohydrates to proteins has long been investigated to alter the physical properties of proteins, for example "neoglycoproteins" from carp myofibrillar protein have been prepared for improved emulsion properties (Saeki, 1997; Sato *et al.*, 2000) in food industry. It has been reported in the synthesis of neoglycoproteins that heating a dry protein-sugar mixture at temperatures greater that 95°C results in glycation but with a significant amount of insoluble products (Boratynski and Roy, 1998; Kanska and Boratynski, 2002).

A significant advantage of nonaqueous chemical modification is that lyophilized proteins are stable at much higher temperatures than in water (Klibanov, 1995) and this property can be used to directly enhance the rate of derivatization. With this knowledge, it seemed worthwhile to investigate whether proteins would glycate under nonaqueous conditions *i.e. in vacuo*, as this presents the possibility of a more rapid reaction with fewer side products due to the absence of water or oxygen. Furthermore, the properties of lyophilized proteins *i.e.* the retention of native-like, biologically active conformational structure, known as "molecular memory" (Yennawar *et al.*, 1995; Mishra *et al.*, 1996) and the "pH memory" can be taken advantage of for the potential to increase the specificity of glycation while retaining biological activity simply by adjusting the pH prior to lyophilization.

Glycation of proteins is an effective method for improving the functional properties of proteins. Synthetic glycoproteins are termed "neoglycoproteins", and many different techniques are available for their preparation (Christopher *et al.*, 1980; Saeki, 1997; Sato *et al.*, 2000). There are an abundant number of methods for the synthesis of such conjugates. As a general rule, the carbohydrate moiety and sometimes also the carrier, have to be chemically converted into their reactive derivatives to facilitate the formation of a covalent bond. In some cases, glycation can be achieved with the use of bifunctional cross-linking reagents or enzymatically by trans-glycosylation (Fan *et al.*, 1995). However, one of the greatest advantages of this method, *in vacuo* glycation of lyophilized proteins, is that it precludes the need for prior derivatization.

This being a novel method for the glycation of proteins, several objectives were set forth as a preliminary assessment. Appropriate conditions for the lyophilization of proteins in the presence of excipient reducing sugars with respect to concentrations and sugar to protein ratio were to be determined. The rate of reaction *in vacuo* of bound sugars with lyophilized proteins and the extent of the *in vacuo* reaction of excipient reducing sugars with proteins lyophilized under various conditions with respect to pH and sugar to protein ratio were also to be determined. As well, the nature of the chemical derivatives formed by reaction of reducing sugars with functional groups in the protein was to be characterized. Finally, the determination of the effect of "pH memory" on the specificity of the glycation of specific functional groups in lyophilized proteins was to be assessed.

Materials and Methods

Proteins and reagents

Ribonuclease A (RNase) (EC 3.1.27.5) from bovine pancreas and human hemoglobin (hHb) were purchased from Sigma Chemical Company.

^{14}C-[U]-D-Glucose (310 mCi/mmol), 50 μCi in 250 μL aqueous solution containing 3% ethanol was purchased from Amersham Pharmacia Biotech. $^{13}C_1$-D-Glucose, 99 atom % ^{13}C and $^{13}C_2$-D-glucose, 99 atom % ^{13}C, were purchased from Sigma Chemical Company. All other chemicals, reagents and solvents used were high purity preparations obtained from commercial sources.

In vacuo glycation of lyophilized proteins

Typically, the protein was added to a glucose solution for a final weight ratio of 10:1 protein: glucose. The pH of the solution was adjusted with 1M HCℓ or NaOH, *via* micro-syringe, to the desired pH. Aliquots of the solution were transferred to borosilicate test tubes, frozen by immersing in liquid nitrogen and lyophilized overnight. The tubes were then narrowed, flame sealed under vacuum (ca. 50-70 mTorr), placed in an oven at 65°C and left to react for a controlled period of time.

Trace *in vacuo* glycation of RNase with ^{14}C-[U]-D-glucose

^{14}C-[U]-D-Glucose (250 nCi) was added to an RNase solution (1mg/mL final concentration). The pH of the solution was adjusted with 1M NaOH *via* micro-syringe, to pH 6.5 or 10.0 and the final volume adjusted to 10 mL with ddH_2O. *In vacuo* glycation was performed as above. Aliquots (50 µL) were transferred to scintillation vials before and after dialysis, scintillation cocktail (5 mL) (Aquasol-2, DuPont) was added and the samples were counted for 10 minutes on a Beckman LS 6500 multi-purpose scintillation counter using the ^{14}C window.

Quantification of *in vacuo* glycation with ^{14}C-[U]-D-glucose

^{14}C-[U]-D-Glucose was added to a protein glucose solution prepared as above for a total of 375 nCi per 10 mg of protein. After reaction, samples were dissolved in 10% (w/v) glucose solution to displace unreacted ^{14}C-[U]-D-glucose and extensively dialyzed (MWCO 3500) against dH_2O (3 x 4L).

Aliquots (50 µL) were transferred to scintillation vials before and after dialysis, scintillation cocktail (5 mL) (Aquasol-2, DuPont) was added and the samples were counted for 10 minutes on a Beckman LS 6500 multi-purpose scintillation counter using the ^{14}C window. Protein amount was normalized by recording absorbance at 280nm on a Pharmacia Biotech Ultraspec 2000 UV-Vis spectrophotometer with dH_2O as blank.

NMR Spectroscopic Analysis

The ^{13}C glycated RNase (50 mg) was extensively dialysed (MWCO 3500) against dH_2O (3 x 4L), lyophilized and dissolved in D_2O (99.9 atom % ^{2}H, Cambridge Isotope Laboratories). The ^{13}C-NMR spectrum was acquired using a Bruker spectrometer operating at 9.4 Tesla (^{13}C, 100.6 MHz) for 600 scans using the DEPT135 pulse sequence (Sanders and Hunter, 1987). Solid state NMR was performed on a Bruker ASX-200 (^{13}C; 50.313 MHz) spectrometer, with a magic angle spinning (MAS) rate of 5 kHz.

Number of scans varied from 2079 to 72065 (overnight acquisition). The co-lyophilized RNase ^{13}C glucose mixture (100 mg) was transferred and packed into the NMR rotor under a nitrogen atmosphere using a glove bag (Aldrich Atmosbag™). To limit sample loss *in vacuo* glycation was performed in the NMR rotor by placing the rotor in a glass tube and sealing it under vacuum. *In vacuo* glycation was performed for 18 hours as described above. Prediction of ^{13}C solid state NMR chemical shifts was performed using the ACD/CNMR Predictor™, Version 4.03 software (Advanced Chemistry Development Inc., Toronto, ON, Canada).

Results and Discussion

Initially, the stability of the protein-glucose lyophilizate was determined by varying protein-sugar ratios for the given reaction conditions. It was qualitatively determined that a weight ratio of 5:1 protein: glucose resulted in the highest sugar content allowed, without the noticeable collapse of the protein-glucose lyophilizate.

Glycation of proteins in water is known to require days to weeks (Neglia *et al.*, 1983). Therefore, a time course using RNase was prepared and carried out in order to compare the rate of glycation *in vacuo* with that of water. Initially, non-aqueous glycation was attempted at two LpH's with a trace amount of ^{14}C glucose (see Figure 4-1).

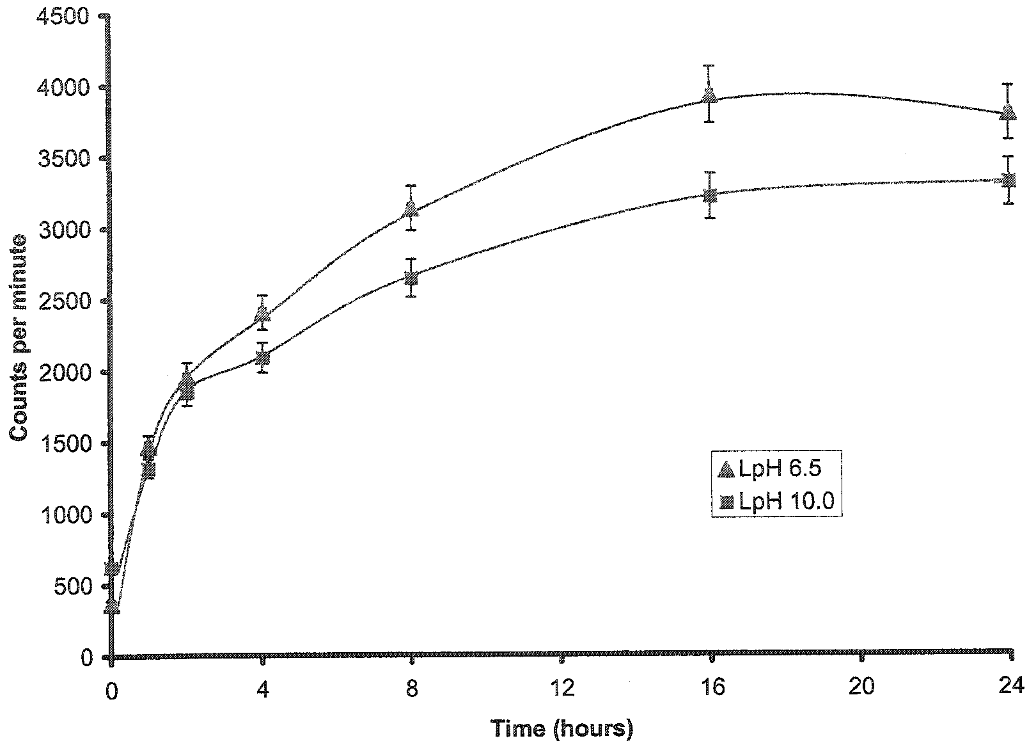


Figure 4-1 Time course and effect of LpH on the *in vacuo* glycation of RNase with a trace of ^{14}C-[U]-D-glucose

The rates of RNase glycation were observed at LpH 6.5 and LpH 10.0. As the results indicate (see Figure 4-1), a great increase in the rate of glycation (hours) was observed as compared to aqueous glycation (days). This greatly increased rate can be explained by the removal of water in the first step when reacting *in vacuo* (see Scheme 4-1) and the increased effective concentration of glucose by lyophilization. From the "pH memory" of lyophilized proteins it was expected that at LpH 6.5 only the α-amino groups would be deprotonated and react thereby giving selective glycation of the α-amino groups. It was observed that both the rate and relative extent of glycation were similar for both LpH. Therefore, it seems that under these conditions there is no apparent pH effect and as a consequence, no preferential selectivity for the α-amino groups over the ε-amino groups as anticipated. These results suggest that the protonated amino groups are readily glycated and therefore the extent of glycation can be easily controlled by limiting the amount of reducing sugar in the protein solution prior to lyophilization. Furthermore, since these results suggest pH independence, glycation can be achieved on the native protein at a pH where the protein's conformation is most stable. The removal of water *via* lyophilization and by placing the lyophilizate *in vacuo* drives the equilibrium towards the Schiff base, based on Le Chatelier's principle. Surprisingly, glycation was complete in just over 16 hours.

The generality of this non-aqueous glycation was also investigated using human hemoglobin (hHb), one of the most thoroughly investigated glycated proteins (Bunn *et al.*, 1978). hHb was lyophilized from a dilute solution of glucose (88.8 mM), with a trace of ^{14}C-[U]-glucose, and lyophilized at pH's 6.5 and 10.0. Although the sugar solution is dilute, it still results in a 70 mole excess of glucose over protein in the lyophilizate.

In sharp contrast to aqueous conditions (Neglia *et al.*, 1983), it was found that under these non-aqueous conditions complete glycation, relative to lysine side chains, could be achieved (see Figure 4-2). All samples showed significant glycation within only 2 hours. Samples at LpH 8.5 were 80% glycated after only 2 hours of heating and the extent of glycation increased to 100% by the 8 hour time point. Samples at LpH 6.5 were just over 40% glycated after only 2 hours and the extent of glycation increased to 100% by the 16 hour time point. Samples at LpH4 were 40% glycated after 2 hours and the extent of glycation increased to over 50% by the 16 hour time point.

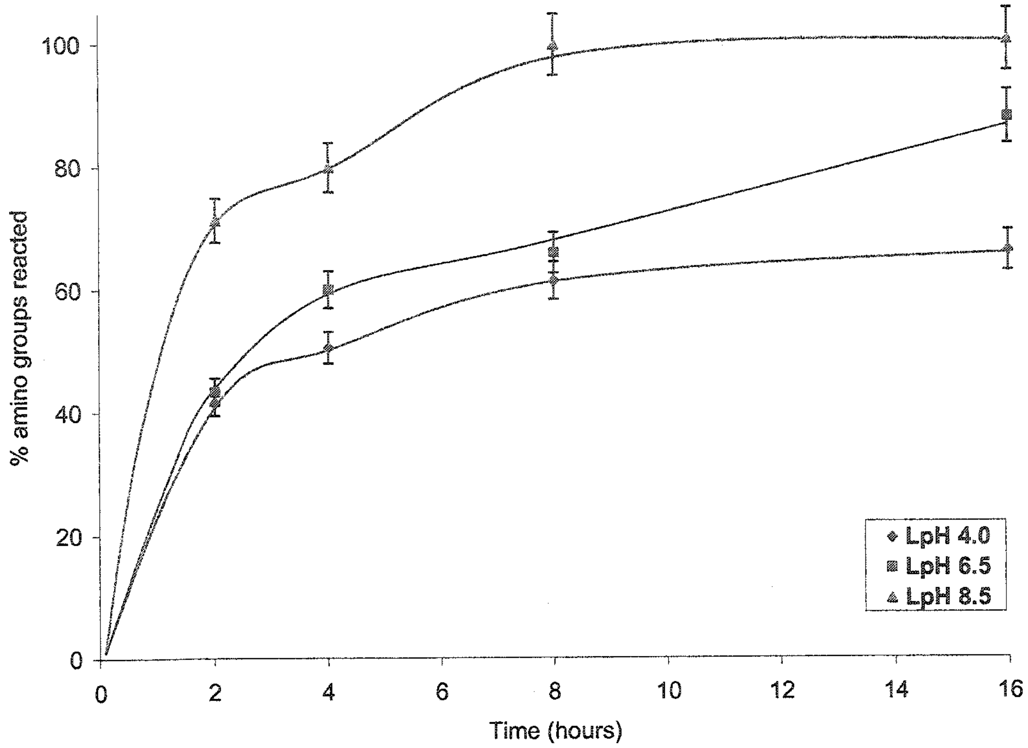


Figure 4-2 Extent of glycation of lyophilized hHb heated *in vacuo* at 65°C, based on 48 possible glycation sites

At LpH 4.0 all amino groups are predicted to be protonated, based on pK_a values of protein amino groups, and also, at LpH 8.5 most of the ε-amino groups are protonated. These results suggest that the classical glycation mechanism in water to form a Schiff base intermediate from the reaction of the un-protonated amino group with the open chain form of the glucose may not occur in the formation of glycoconjugate in the lyophilized state.

The acid-base solution stability of the glycation adducts was investigated at three different pH's and two different temperatures (see Figure 4-3 and Figure 4-4). The ketoamine products under aqueous conditions both seem to be the most stable under neutral conditions and degrade slowly at acidic and alkaline pH values. It also appears that higher temperatures consistently resulted in higher lability of the ketoamine. These results can be explained by either a dissociation of glucose from the protein in the presence of water as in theory an equilibrium can be re-established (see Scheme 1) or by some extent of degradation of the protein itself. If an equilibrium is re-established, the loss of glucose would be expected to be greater than that observed as the glucose is not in excess.

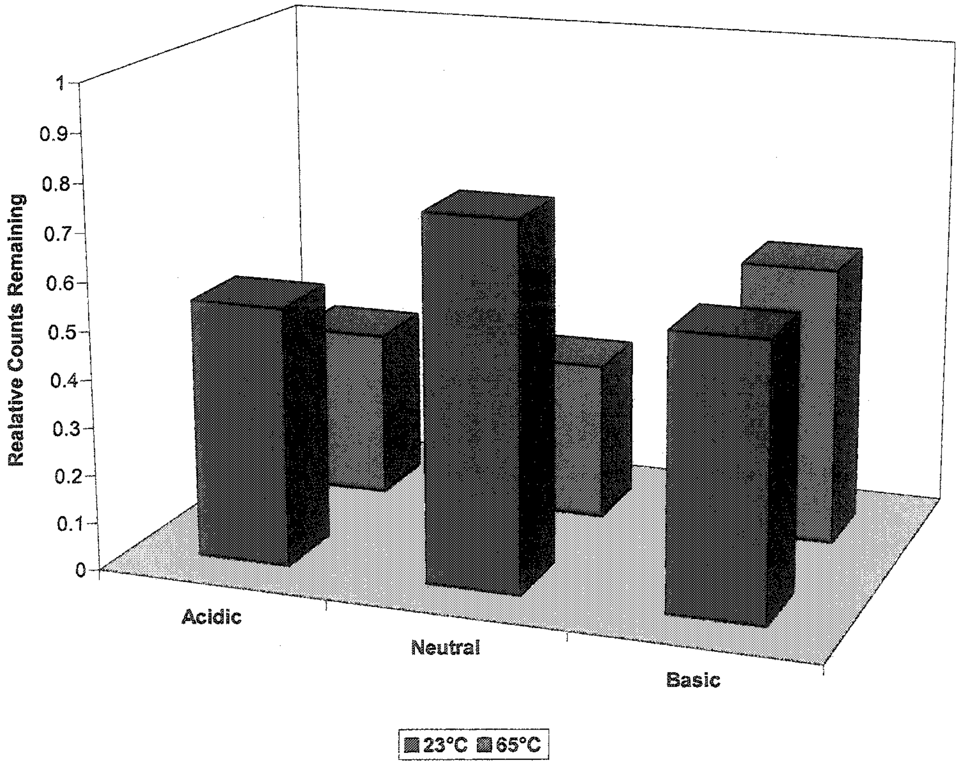


Figure 4-3 Acid-base lability of LpH 6.5 glycated RNase. The ^{14}C glycated RNase (LpH 6.5) was incubated in phosphate buffer (200 mM) at pH 2.0 (acid), pH 7.0 (neutral), and pH 10.5 (base). One set was incubated at 23°C and the other at 65°C and both were left to hydrolyze for a period of 24 hours

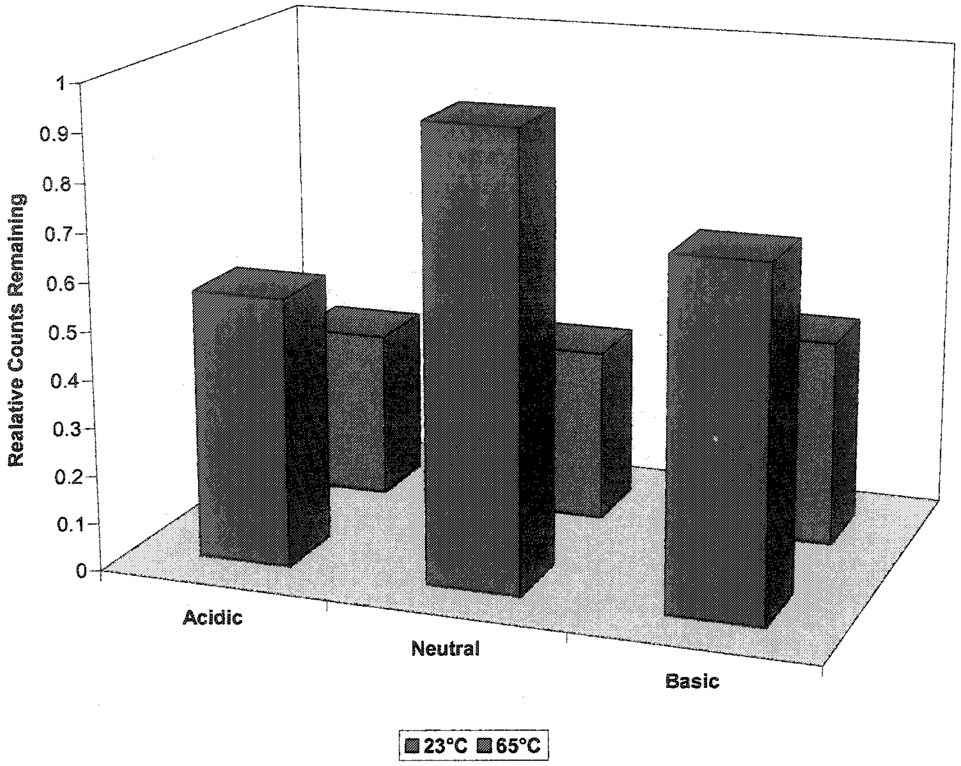


Figure 4-4 Acid-base lability of LpH 10.0 glycated RNase. The ^{14}C glycated RNase (LpH 10) was incubated in phosphate buffer (200 mM) at pH 2.0 (acid), pH 7.0 (neutral), and pH 10.5 (base). One set was incubated at 23°C and the other at 65°C and both were left to hydrolyze for a period of 24 hours

RNase was glycated *in vacuo* with $^{13}C_1$-D-glucose in an attempt to gain insight as to the type(s) of modification afforded. RNase was chosen as the model protein as its glycation in aqueous solution has already been studied and characterized by ^{13}C solution NMR (Neglia *et al.*, 1983; Neglia *et al.*, 1985). Solution ^{13}C-NMR of the glycated lyophilized RNase with $^{13}C_1$-D-glucose revealed that only one product is present after dialysis. The single resonance peak, located at 53.2 ppm (see Figure 4-5), corresponds to the cyclic form of the protonated ketoamine adduct (Neglia *et al.*, 1983). This result clearly demonstrates the highly selective nature of this novel approach for protein glycation.

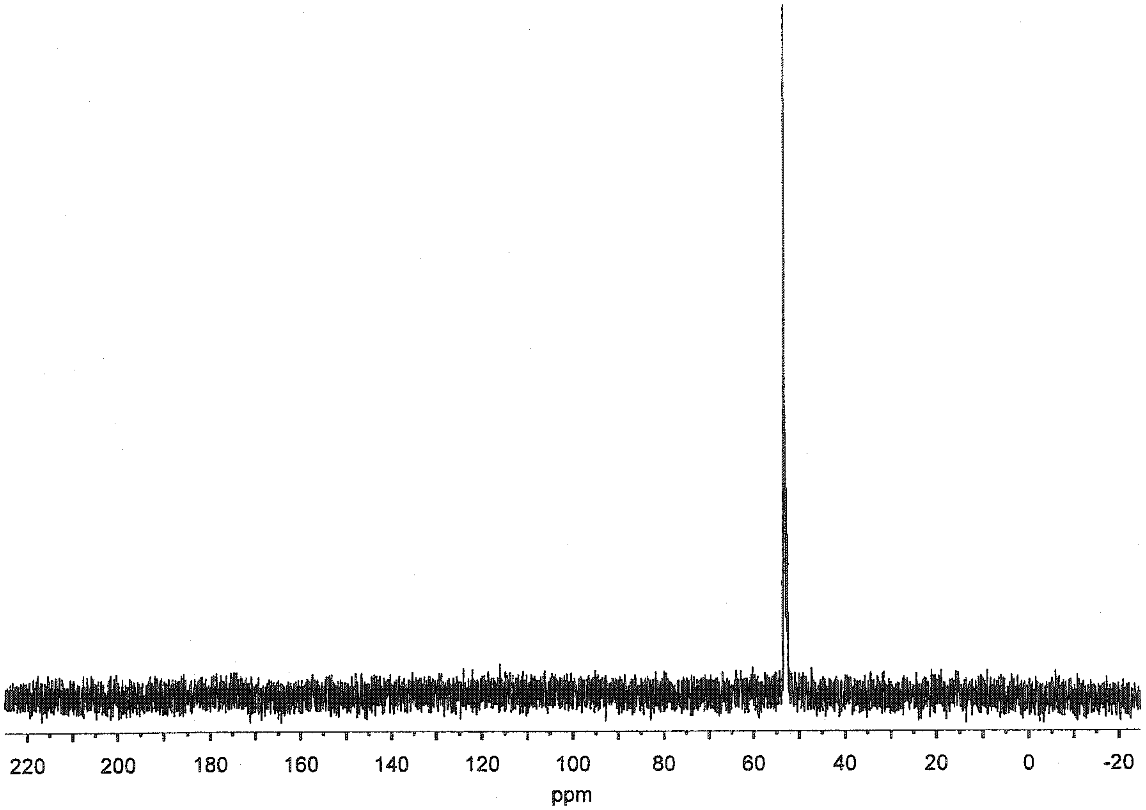
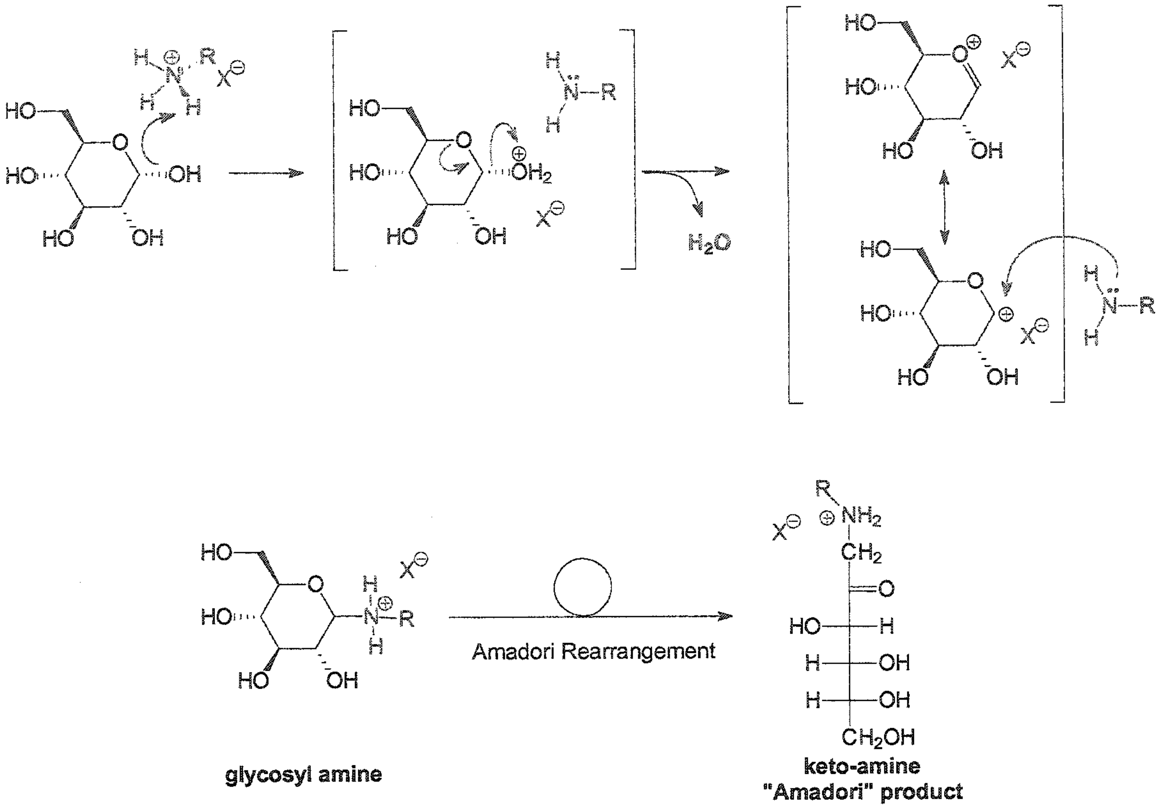


Figure 4-5 Solution ^{13}C-NMR spectrum of *in vacuo* glycated RNase with $^{13}C_1$-D-glucose

The exact mechanism of *in vacuo* glycation is not clear. In the aqueous reaction (see Scheme 4-1), the open chain form of the reducing sugar is needed to form the Schiff base. The accepted mechanism for glycation in water involves the nucleophilic attack of a protein's amino groups on the open chain form of the reducing sugar, forming a Schiff base intermediate, which then undergoes Amadori rearrangement to the keto-amine product (Neglia *et al.*, 1983). Therefore, very little glycation was expected below LpH 7 since the bulk of the amino groups of the protein are in the protonated state. Surprisingly, the results from the *in vacuo* glycation strongly suggest that the mechanism does not necessarily involve the formation of a Schiff base and that glycation can occur with the protonated amino groups.

In solution, ketoses are known to undergo intramolecular nucleophilic attack by the C-5 or C-6 hydroxyl groups on the carbonyl carbon, producing five- or six-membered cyclic hemiketals. This process converts an achiral carbonyl carbon into a chiral hemiketal and produces two isomers that are termed anomers, designated α and β. Therefore a solution of a ketose at equilibrium contains a mixture of α and β furanoses and a mixture of α and β pyranoses in addition to the acyclic form (El Khadem, 1988). The amount of the acyclic form of the reducing sugar (~0.4%) is a rate-limiting step in the aqueous glycation reaction (Roy *et al.*, 1984). The acyclic to cyclic ratio of the reducing sugar is not expected to be different in the lyophilizate, therefore two possible mechanisms are proposed for *in vacuo* glycation. The first proposed mechanism is that the reaction proceeds *via* the cyclic form of the reducing sugar (see Scheme 4-2) with a proton transfer analogous to one of the reaction mechanisms proposed for the *in vacuo* reaction of iodomethane with peptides (*vide supra*), followed by the loss of water vapour to the vacuum. The second is that the cyclic anomers are in equilibrium through the acyclic form, as in solution, and as the acyclic conformation is formed it reacts (see Scheme 4-1).

The proposed glycation mechanism may be substantiated by an increased effective concentration of glucose and its close proximity to the protonated amino groups in the lyophilized state. This phenomenon appears to be protein dependent as observed when comparing RNase to hemoglobin. RNase in solution is found as a monomer whereas hemoglobin is found as a tetramer. This structural difference may directly affect the accessibility of the amino groups to the reducing sugar and the extent of glycation. Studies on other proteins are required to clarify this point.



R = protein attached via α or ε-amino group
X = anion (eg. protein carboxylate or other from lyophilized solution)

Scheme 4-2 Hypothesized *in vacuo* glycation of proteins with D-glucose

Cross polarization magic angle spinning (CPMAS) solid state NMR was also used in hopes to further elucidate the reaction mechanism of *in vacuo* glycation. The solid state ^{13}C-NMR spectrum of RNase and $^{13}C_1$-glucose co-lyophilized from an aqueous solution at pH 8.0 was recorded prior to heating *in vacuo* (see Figure 4-6). The resonances attributed to the natural abundance ^{13}C from the protein are clearly visible. Aliphatic carbons are in the region of 10 to 70 ppm and the carbonyl carbons make up the band centered at 172.7 ppm. The strongest peaks arise from the α and β anomeric positions of the $^{13}C_1$ of glucose at 97.2 and 93.1 ppm, respectively. From this spectrum it is clear that no appreciable amount of glycation occurs during the lyophilization process. The two peaks from the anomeric carbon are narrower due to the fact that this is a small molecule.

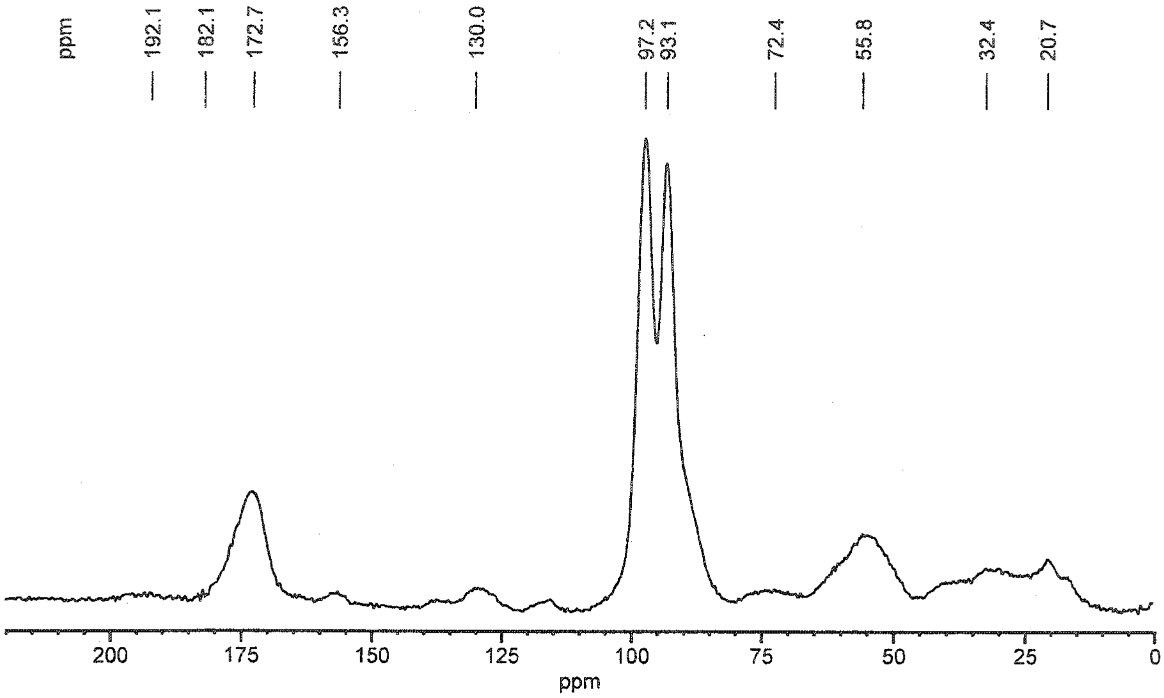


Figure 4-6 CPMAS ^{13}C-NMR spectrum of RNase co-lyophilized with $^{13}C_1$-D-glucose

The solid state ^{13}C-NMR spectrum of glycated RNase with $^{13}C_1$-D-glucose is shown in Figure 4-7. The resonances of the α and β anomeric carbon nearly disappear with the emergence of two peaks; one at 54.8 ppm, which would correspond to the predicted protonated ketoamine and one at 25.0 ppm, with narrower line width suggesting a small molecule. The presence of a Schiff base is not evident, which implies it to be an intermediate formed after the initial loss of water, similar to the aqueous reaction. A glycosyl amine with a predicted chemical shift of ca. 89 ppm (M.C. King, PhD thesis) is also not observed.

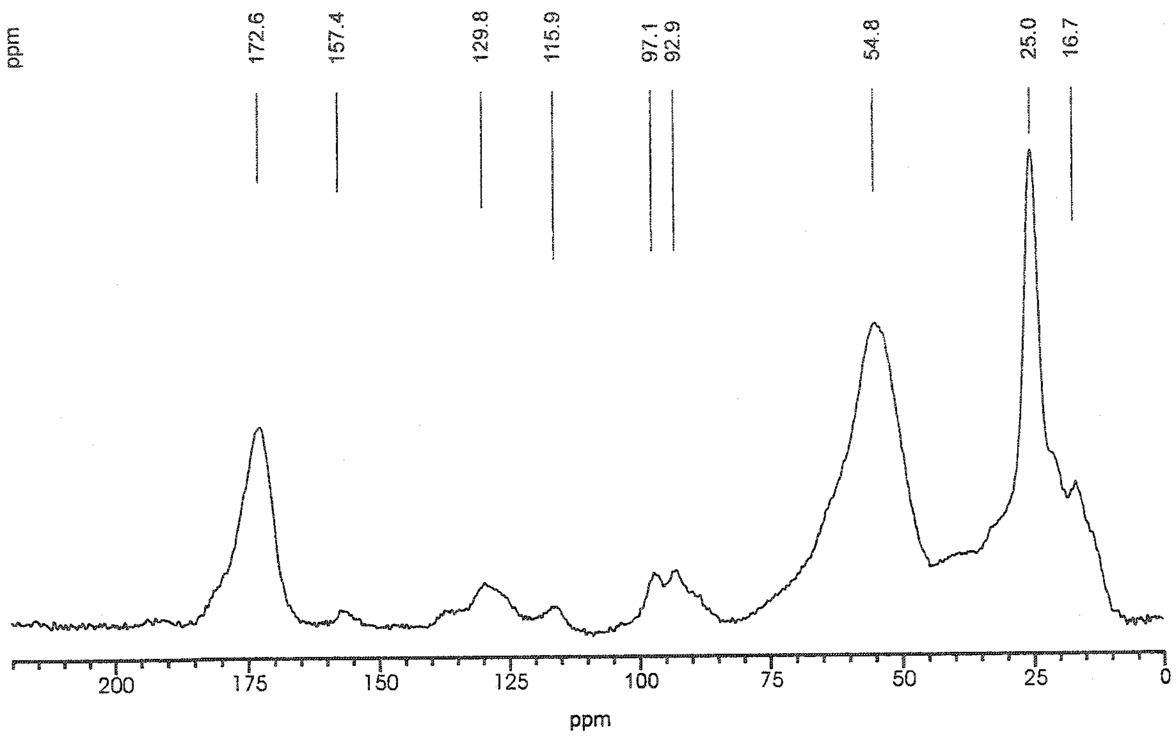


Figure 4-7 CPMAS ^{13}C-NMR spectrum of *in vacuo* glycated RNase with $^{13}C_1$-D-glucose

To further investigate the nature of the unknown resonances a dipolar dephasing (DD) experiment was employed to suppress the signals of protonated carbons. The DD experiment uses differences in the strength of C-H dipolar coupling to enable the distinction between protonated and non-protonated carbons, and molecularly mobile and rigid carbons. The strongest spin interaction for most carbons is dipolar coupling to protons. The ^{13}C-^{1}H dipolar interaction has a $1/r^3$ dependence, where r is the internuclear distance; therefore, the coupling is much greater for carbons bonded to protons ($r \approx 1.1$ Å) compared with those carbons with protons only on the neighbour carbons ($r > 2.0$ Å). In a DD experiment, the high power proton decoupler is turned off for a short period (the dephasing delay) between polarization and detection, during which time ^{13}C signal is lost through T_2 broadening. The rate of signal loss is dependent on the strength of C-H dipolar coupling, which is primarily determined by the proximity of nearest proton neighbours. Carbons with directly attached protons generally dephase rapidly, with little signal left after 40-60 μs, whereas non-protonated carbons dephase more slowly (Opella and Frey, 1979; Opella *et al.*, 1979). This can be used to differentiate between the two types of carbons although the precise number of attached protons cannot be identified from the DD ^{13}C solid-state spectra. A complication can arise with this experiment from carbons possessing high molecular motion. Methyl groups dephase slowly despite the presence of three attached protons due to rapid rotation which greatly diminishes the strength of coupling.

Figure 4-8 shows the CPMAS-DD ^{13}C-NMR spectrum of *in vacuo* glycated RNase with $^{13}C_1$-D-glucose.

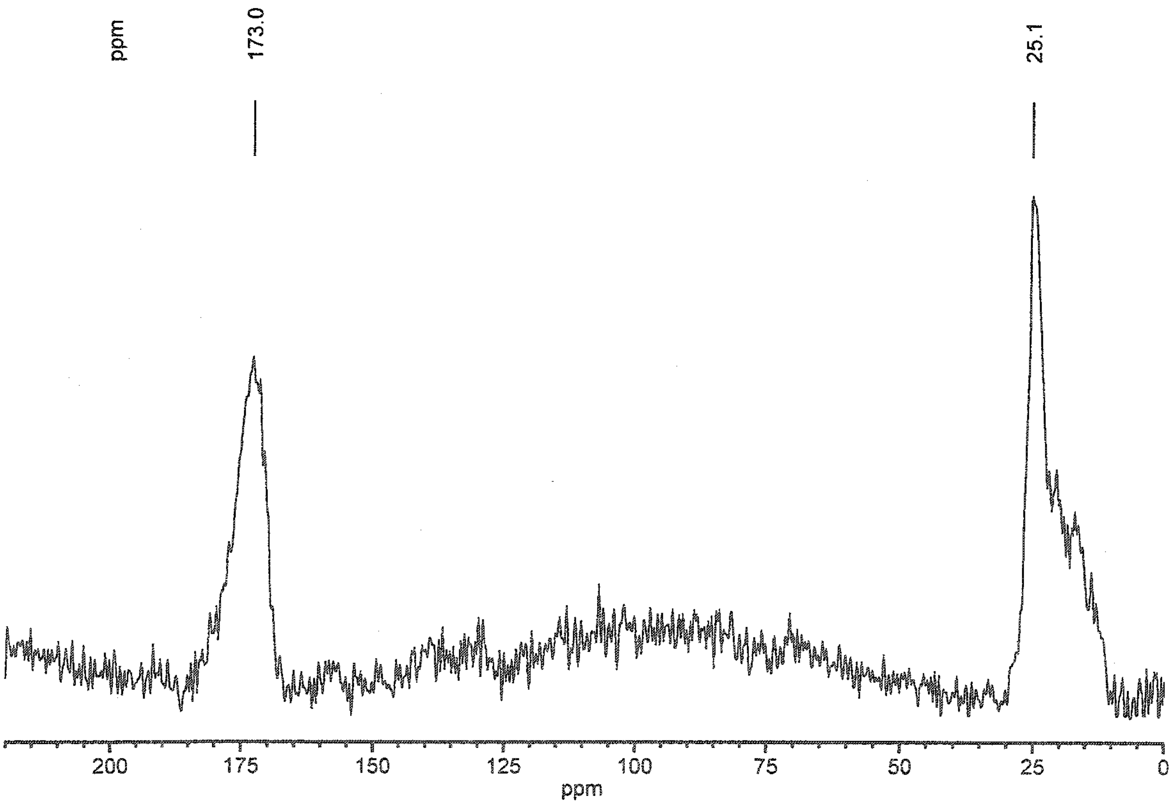


Figure 4-8 CPMAS-DD ^{13}C-NMR spectrum of *in vacuo* glycated RNase with $^{13}C_1$-D-glucose

The narrow peak at 25.1 ppm remains, along with some resonances below 25 ppm shouldering this peak and resonances centred at 173.0 ppm corresponding to the natural abundance ^{13}C quaternary carbonyl carbons of the protein.

The peak at 54.8 ppm is predicted to be from the protonated ketoamine, based on its chemical shift (Yaylayan and Huyghues-Despointes, 1994). This peak totally disappears during DD, suggesting that the carbon giving rise to this resonance has protons attached to it. Indeed, the C-1 carbon of the ketoamine has two protons. The peak at 25.1 ppm cannot be a quaternary carbon as the chemical shift is too high and actually falls into the aliphatic resonance region and therefore is suspected to be a mobile methyl carbon. The fact that it is a narrow peak also suggests that it may be a small molecule. The origin of the peak at 25.1 ppm is not clear and a similar experiment using $^{13}C_2$-D-glucose was performed to perhaps elucidate the origin of this peak. The solid state ^{13}C-NMR spectrum of glycated RNase with $^{13}C_2$-D-glucose is shown in Figure 4-9.

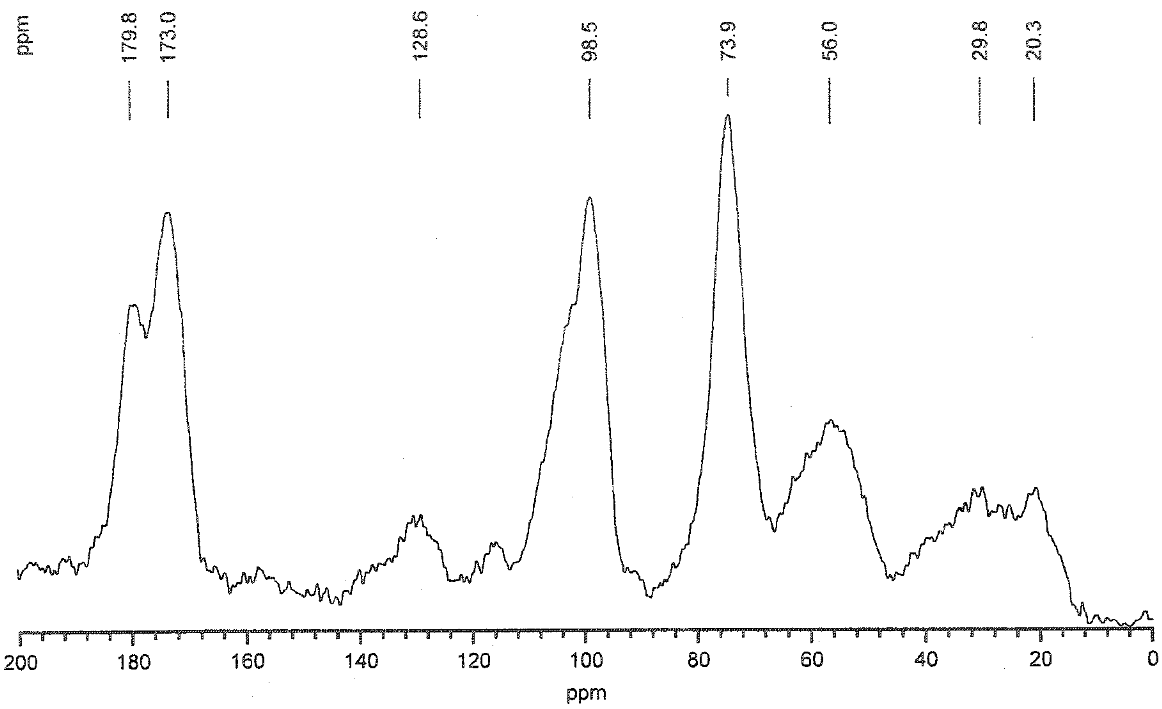


Figure 4-9 CPMAS ^{13}C-NMR spectrum of RNase glycated *in vacuo* with $^{13}C_2$-D-glucose

Unfortunately the spectrum prior to heating *in vacuo* was not recorded but it is expected that, similar to the $^{13}C_1$ experiment, no observable glycation occurred during the lyophilization step and all of the glucose reacts with the protein. The natural abundance ^{13}C resonances from the protein are clearly seen, as previously observed, with the addition of three strong peaks at 179.8, 98.5 and 73.9 ppm. The peak at 98.5 ppm would correspond with the hemiacetal C-2 carbon of the ketoamine as it remains in the CPMAS-DD ^{13}C-NMR spectrum (see Figure 4-10). The peak at ca.180 ppm also remains suggesting that it is possesses no protons whereas the peak at 73.9 ppm totally disappears, again suggesting a carbon with protons. The origin of these resonance signals is not clear but they may arise from the intermediate(s) or side products described earlier, as their chemical shifts do not correspond with that of the ketoamine product.

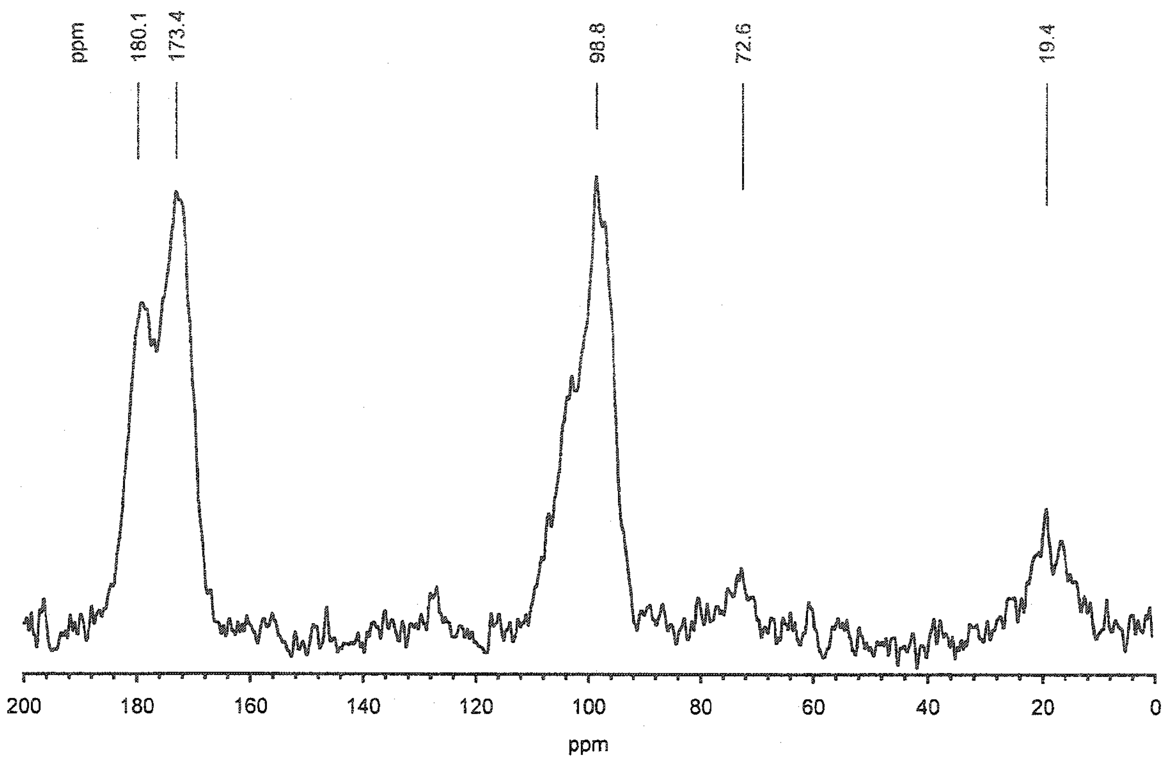
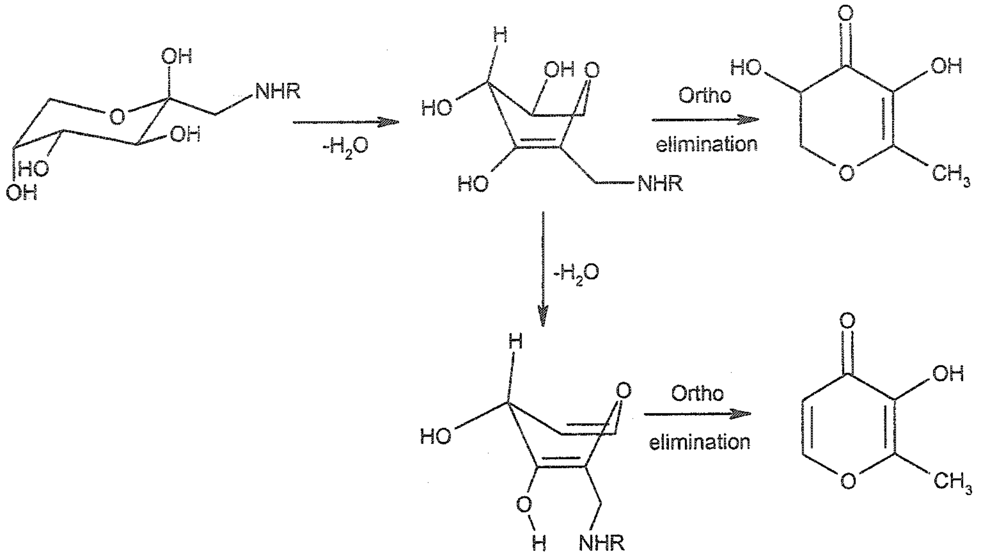


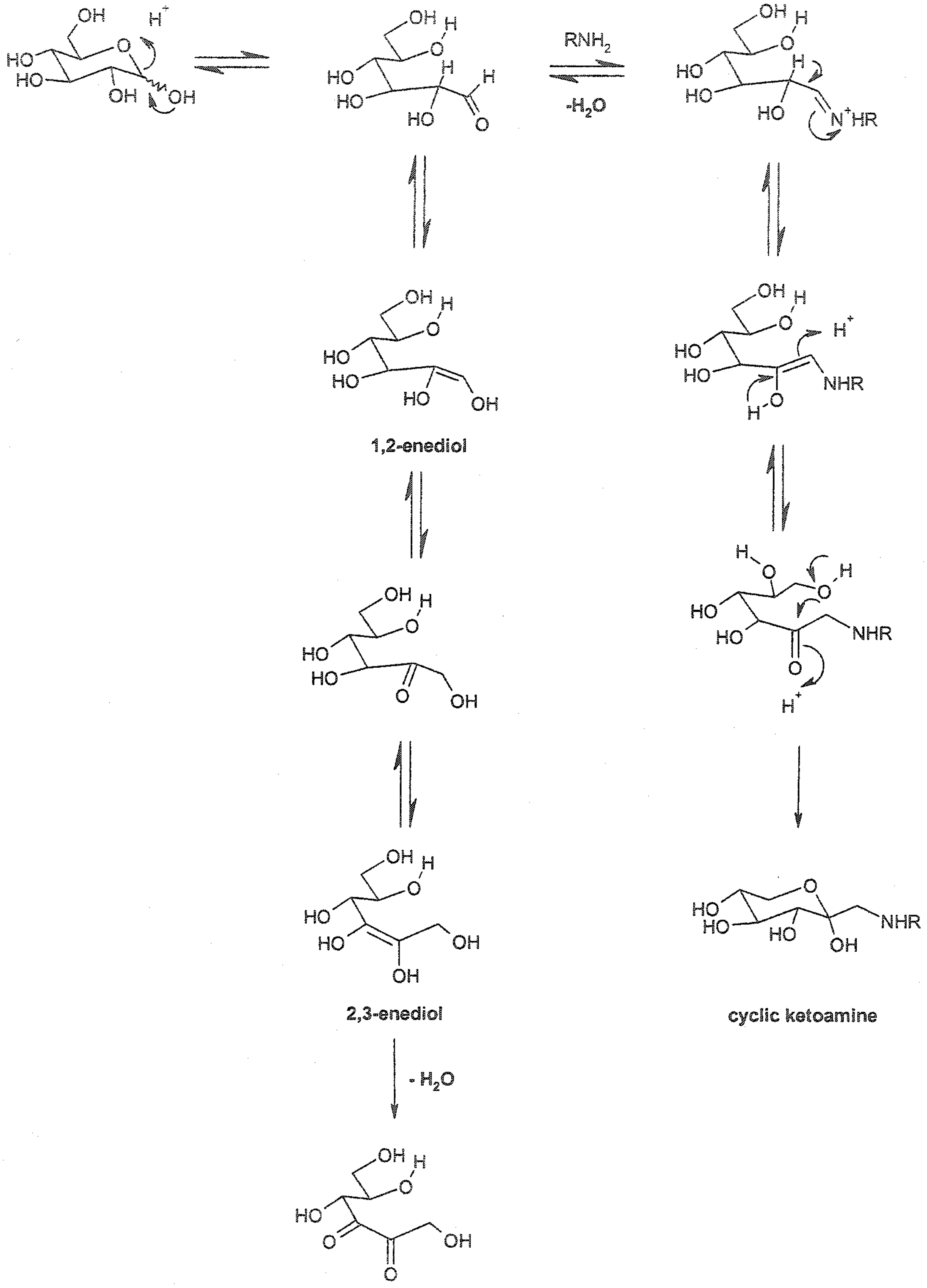
Figure 4-10 CPMAS-DD ^{13}C-NMR spectrum of RNase glycated *in vacuo* with $^{13}C_2$-D-glucose

Reinvestigation of the $^{13}C_1$-glucose solid state NMR spectrum (Figure 4-7) a peak at the carbonyl region (ca. 173ppm) stated to originate from the protein's carbonyl carbons may also harbour a carbonyl signal from another side product as it has gained some intensity as compared to the un-heated control. This suggests the possibility of one or two small side products or unstable intermediates formed during *in vacuo* glycation as the solution NMR spectrum (Figure 4-5) clearly shows that only the ketoamine remains after dialysis. The exact nature of these side products cannot be determined from the spectra. The methyl group suspected in the DD spectrum could originate from a secondary dehydration step, perhaps facilitated by the *in vacuo* environment. The side product, as described by Yaylayan and Huyghues-Despointe, is proposed to originate from the ortho elimination of the cyclic form of the Amadori product in the presence of an acid is presented in Scheme 4-3 (Yaylayan and Huyghues-Despointes, 1994).



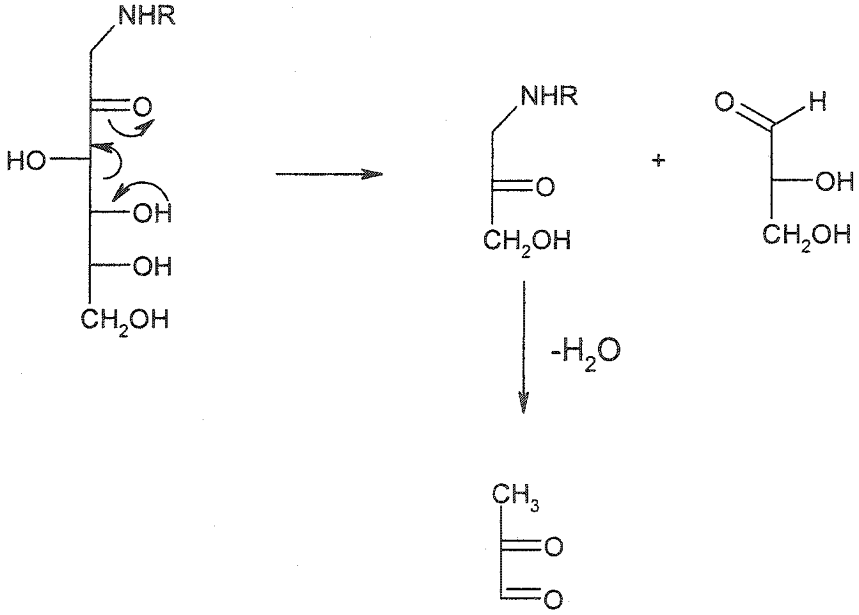
Scheme 4-3 Proposed ortho-elimination from the ketoamine as a potential product observed in the solid state NMR

This could be explained in terms of the lyophilization process increasing the effective concentration of glucose and its close proximity to a proton donor or acceptor. If this is a possibility the enolization of glucose itself cannot be ruled out. In light of the results a complete glycation scheme is presented (see Scheme 4-4) with the enolization of glucose shown as a possible side reaction during *in vacuo* glycation. The proton in the initial step may come from the protonated ammonium group of the protein or some other acidic source.



Scheme 4-4 Complete glycation scheme with the enolization of glucose shown as a possible side reaction during *in vacuo* glycation

Another possible side reaction occurring to give resonances similar in chemical shift to those observed would be the retro-aldol reaction of the ketoamine followed by the *β*-elimination of water, again perhaps facilitated by the vacuum, to give the diketone products as shown in Scheme 4-5. Unfortunately diketones are quite reactive towards amino groups and especially reactive towards quanidino groups (*i.e.* arginine) (Biemel *et al.*, 2002) and may elicit the formation of other AGE's. These types of AGEs have been shown to be involved in the creation of cross links through arginine residues (Munch *et al.*, 1999).



Scheme 4-5 Retro-aldol and β-elimination of water as another potential route to the product observed in the solid state NMR from the Amadori product

One can speculate as to the nature of the intermediates or side products formed but the fact remains that the *in vacuo* glycation proceeds by the protonated amino groups of proteins. Although the glycation of a protonated amino group is unusual it is not unprecedented. An example of glycation occurring in the dry state via the protonated amine has been found with the case of Wirth *et al.* It was found that lactose and a secondary amine; fluoxetine·HCℓ (Prozac™), reacted to form coloured pigments via the Maillard reaction (Wirth *et al.*, 1998). Prior to this example it was thought that the hydrochloride salts of amines could not react with reducing sugars. The reaction was explained to perhaps occur with the initial loss of HCℓ gas from the hydrochloride salt.

The results obtained describe a novel approach for the synthesis of neoglycoconjugates. The results show that reducing sugars covalently bind to proteins after they are lyophilized together to dryness and then left to react *in vacuo*. Optimization of conditions resulted in a protocol for the simple production of highly, yet selectively, substituted glycoproteins. This method allows for the direct attachment of carbohydrate moieties to the protein without any prior modification or chemical activation.

The realization of non-aqueous glycation owes much credit to the process of lyophilization. This method imparts two very important and advantageous characteristics on the given system. First, much higher temperatures can now be used without risk of denaturing the protein (Klibanov, 1997). Second, there is an increase in the effective concentration of the solution. This increase in effective concentration brings the molecules closer together and believed to play a leading role in the high incorporation of glucose that was achieved.

In summary, *in vacuo* glycation proceeds much more rapidly than in water via the protonated amino groups of proteins to give one predominant adduct: the cyclic Amadori ketoamine. This method shows much promise for the selective glycation of proteins to modify the physical properties and in turn function of the protein.

Chapter 5 Future endeavours

The unique property of the trimethylammonium derivative of amino groups of peptides and the significant increased sensitivity in MALDI MS makes this an attractive tool for characterization purposes. This permanent positive charge can be used for the isolation of N-terminal peptides by a "diagonal" electrophoretic method (similar to the C-terminal method described herein) with the use of MALDI MS, either for quantification or as a means of peptide mapping for comparative studies. This is currently being pursued in Dr. Kaplan's laboratory during the writing of this thesis by a PhD. student Pham Van Thong

In vacuo glycation of other amino containing molecules, such as peptides, solid supports and dendrimers by simple co-lyophilization with glucose should also be investigated. The application of other reducing sugars and molecules containing one or more reducing ends should be explored as a potential synthetic route to immobilization of proteins or protein cross linking. This has been investigated by Mary King, PhD. and submitted as a patent.

References

Allen, G. (1981). Sequencing of Proteins and Peptides. North-Holland, Elsevier.

Barman, T. E. (1969). "Enzyme handbook." Berlin, Heidelberg, New York, Springer. **II:** 620-621.

Beardsley, R. L., J. A. Karty and J. P. Reilly (2000). "Enhancing the intensities of lysine-terminated tryptic peptide ions in matrix-assisted laser desorption/ionization mass spectrometry." *Rapid Commun Mass Spectrom* **14**(23): 2147-53.

Bell, L. N., M. J. Hageman and L. M. Muraoka (1995). "Thermally induced denaturation of lyophilized bovine somatotropin and lysozyme as impacted by moisture and excipients." *J Pharm Sci* **84**(6): 707-12.

Bergman, T., E. Cederlund and H. Jornvall (2001). "Chemical C-terminal protein sequence analysis: improved sensitivity, length of degradation, proline passage, and combination with edman degradation." *Anal Biochem* **290**(1): 74-82.

Bertozzi, C. R. and L. L. Kiessling (2001). "Chemical glycobiology." *Science* **291**(5512): 2357-64.

Biemel, K. M., D. A. Friedl and M. O. Lederer (2002). "Identification and quantification of major maillard cross-links in human serum albumin and lens protein. Evidence for glucosepane as the dominant compound." *J Biol Chem* **277**(28): 24907-15.

Billeci, T. M. and J. T. Stults (1993). "Tryptic mapping of recombinant proteins by matrix-assisted laser desorption/ionization mass spectrometry." *Anal Chem* **65**(13): 1709-16.

Blow, D. M., J. J. Birktoft and B. S. Hartley (1969). "Role of a buried acid group in the mechanism of action of chymotrypsin." *Nature* **221**(178): 337-40.

Bonetto, V., A. C. Bergman, H. Jornvall and R. Sillard (1997). "C-terminal sequence analysis of peptides and proteins using carboxypeptidases and mass spectrometry after derivatization of Lys and Cys residues." *Anal Chem* **69**(7): 1315-9.

Bonetto, V., A. C. Bergman, H. Jornvall and R. Sillard (1997). "C-terminal sequence determination of modified peptides by MALDI MS." *J Protein Chem* **16**(5): 371-4.

Boratynski, J. and R. Roy (1998). "High temperature conjugation of proteins with carbohydrates." *Glycoconj J* **15**(2): 131-8.

Bornsen, K. O. (2000). "Influence of salts, buffers, detergents, solvents, and matrices on MALDI-MS protein analysis in complex mixtures." *in* Mass spectrometry of proteins and peptides. J. R. Chapman. Totowa, N.J., Humana Press: 387-404.

Boyle, W. J., P. van der Geer and T. Hunter (1991). "Phosphopeptide mapping and phosphoamino acid analysis by two-dimensional separation on thin-layer cellulose plates." *Methods Enzymol* **201**: 110-49.

Braco, L., K. Dabulis and A. M. Klibanov (1990). "Production of abiotic receptors by molecular imprinting of proteins." *Proc Natl Acad Sci U S A* **87**(1): 274-7.

Brancia, F. L., S. G. Oliver and S. J. Gaskell (2000). "Improved matrix-assisted laser desorption/ionization mass spectrometric analysis of tryptic hydrolysates of proteins following guanidination of lysine-containing peptides." *Rapid Commun Mass Spectrom* **14**(21): 2070-3.

Broos, J., A. Visser, J. F. J. Engbersen, W. Verboom, A. Vanhoek and D. N. Reinhoudt (1995). "Flexibility of enzymes suspended in organic solvents probed by time-resolved fluorescence anisotropy - evidence that enzyme activity and enantioselectivity are directly related to enzyme flexibility." *J Am Chem Soc* **117**(51): 12657-12663.

Bucala, R., Z. Makita, G. Vega, S. Grundy, T. Koschinsky, A. Cerami and H. Vlassara (1994). "Modification of low density lipoprotein by advanced glycation end products contributes to the dyslipidemia of diabetes and renal insufficiency." *Proc Natl Acad Sci U S A* **91**(20): 9441-5.

Bunn, H. F., K. H. Gabbay and P. M. Gallop (1978). "The glycosylation of hemoglobin: relevance to diabetes mellitus." *Science* **200**(4337): 21-7.

Burke, P. A., R. G. Griffin and A. M. Klibanov (1992). "Solid-state NMR assessment of enzyme active center structure under nonaqueous conditions." *J Biol Chem* **267**(28): 20057-64.

Byrn, S. R., W. Xu and A. W. Newman (2001). "Chemical reactivity in solid-state pharmaceuticals: formulation implications." *Adv Drug Deliv Rev* **48**(1): 115-36.

Chapman, J. R. (2000). Mass spectrometry of proteins and peptides. Totowa, N.J., Humana Press.

Chen, C. S. and C. J. Sih (1989). "General aspects and optimization of enantioselective biocatalysis in organic solvents. The use of lipases." *Angewandte Chemie International Edition in English* **28**: 695-707.

Christopher, P., P. Stowell and Y. C. Lee (1980). "Neoglycoproteins. The preparation and application of synthetic glycoproteins." *Adv Carbohydrate Chem Biochem* **37**: 225-281.

Cleland, J. L., R. S. Langer and American Chemical Society. Division of Biochemical Technology. (1994). Formulation and delivery of proteins and peptides. Washington, DC, American Chemical Society.

Cohen, G. H., E. W. Silverton and D. R. Davies (1981). "Refined crystal structure of gamma-chymotrypsin at 1.9 A resolution. Comparison with other pancreatic serine proteases." *J Mol Biol* **148**(4): 449-79.

Cohen, S. L. and B. T. Chait (1997). "Mass spectrometry of whole proteins eluted from sodium dodecyl sulfate-polyacrylamide gel electrophoresis gels." *Anal Biochem* **247**(2): 257-67.

Colaco, C. A., M. D. Ledesma, C. R. Harrington and J. Avila (1996). "The role of the Maillard reaction in other pathologies: Alzheimer's disease." *Nephrol Dial Transplant* **11 Suppl 5**: 7-12.

Costantino, H. R., K. Griebenow, P. Mishra, R. Langer and A. M. Klibanov (1995). "Fourier-transform infrared spectroscopic investigation of protein stability in the lyophilized form." *Biochim Biophys Acta* **1253**(1): 69-74.

Creighton, T. E. (1993). Proteins : structure and molecular properties. New York, W.H. Freeman.

Cruickshank, W. H. and H. Kaplan (1972). "A competitive labelling method for determining the ionization constants and reactivity of individual histidine residues in proteins. The histidines of alpha-chymotrypsin." *Biochem J* **130**(4): 1125-31.

Cruickshank, W. H. and H. Kaplan (1974). "Properties of the histidines of chymotrypsinogen: comparison with alpha-chymotrypsin." *J Mol Biol* **83**(2): 267-74.

Dabulis, K. and A. M. Klibanov (1993). "Dramatic enhancement of enzymatic activity in organic solvents by lyoprotectants." *Biotechnol Bioeng* **41**: 566-571.

Darbre, A. (1986). Practical protein chemistry : a handbook. Chichester [Sussex] ; New York, Wiley.

Duggleby, R. G. and H. Kaplan (1975). "A general method for the determination of the carboxyl-terminal sequence of proteins." *Anal Biochem* **65**(1-2): 346-54.

Edman, P. and G. Begg (1967). "A protein sequenator." *Eur J Biochem* **1**(1): 80-91.

Ehring, H., S. Stromberg, A. Tjernberg and B. Noren (1997). "Matrix-assisted laser desorption/ionization mass spectrometry of proteins extracted directly from sodium dodecyl sulphate-polyacrylamide gels." *Rapid Commun Mass Spectrom* **11**(17): 1867-73.

El Khadem, H. S. (1988). Carbohydrate Chemistry. San Diego, CA., Academic Press.

Fan, J. Q., M. S. Quesenberry, K. Takegawa, S. Iwahara, A. Kondo, I. Kato and Y. C. Lee (1995). "Synthesis of neoglycoconjugates by transglycosylation with Arthrobacter protophormiae endo-beta-N-acetylglucosaminidase. Demonstration of a macro-cluster effect for mannose-binding proteins." *J Biol Chem* **270**(30): 17730-5.

Fenn, J. B., M. Mann, C. K. Meng, S. F. Wong and C. M. Whitehouse (1989). "Electrospray ionization for the mass spectrometry of large biomolecules." *Science* **246**: 64-71.

Fersht, A. (1985). Enzyme structure and mechanism. New York, W.H. Freeman.

Furka, A., G. Dibo, J. Kovacs and F. Sebestyen (1983). "An improved method for isolation of the C-terminal fragment of proteins." *Anal Biochem* **129**(1): 14-21.

Furka, A., F. Sebestyen and T. Karacsonyi (1970). "New method for isolation of the C-terminal fragment of peptides and proteins." *FEBS Lett* **6**(1): 34-36.

Gevaert, K., H. De Mol, J. L. Verschelde, J. Van Damme, S. De Boeck and J. Vandekerckhove (1997). "Novel techniques for identification and characterization of proteins loaded on gels in femtomole amounts." *J Protein Chem* **16**(5): 335-42.

Godovac-Zimmermann, J. and L. R. Brown (2001). "Perspectives for mass spectrometry and functional proteomics." *Mass Spectrom Rev* **20**(1): 1-57.

Goodlett, D. R., A. Keller, J. D. Watts, R. Newitt, E. C. Yi, S. Purvine, J. K. Eng, P. von Haller, R. Aebersold and E. Kolker (2001). "Differential stable isotope labeling of peptides for quantitation and de novo sequence derivation." *Rapid Commun Mass Spectrom* **15**(14): 1214-21.

Griebenow, K. and A. M. Klibanov (1997). "Can conformational changes be responsible for solvent and excipient effects on the catalytic behavior of subtilisin carlsberg in organic solvents." *Biotechnol Bioeng* **53**(4): 351-362.

Griffin, T. J., S. P. Gygi, B. Rist, R. Aebersold, A. Loboda, A. Jilkine, W. Ens and K. G. Standing (2001). "Quantitative proteomic analysis using a MALDI quadrupole time-of-flight mass spectrometer." *Anal Chem* **73**(5): 978-86.

Gusev, A. I., W. R. Wilkinson, A. Proctor and D. M. Hercules (1995). "Improvement of signal reproducibility and matrix/comatrix effects in MALDI analysis." *Anal Chem* **67**(6): 1034-1041.

Gygi, S. P., B. Rist and R. Aebersold (2000). "Measuring gene expression by quantitative proteome analysis [Review]." *Curr Opin Biotechnol* **11**(4): 396-401.

Gygi, S. P., B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb and R. Aebersold (1999). "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags." *Nat Biotechnol* **17**(10): 994-9.

Hale, J. E., J. P. Butler, M. D. Knierman and G. W. Becker (2000). "Increased sensitivity of tryptic peptide detection by MALDI-TOF mass spectrometry is achieved by conversion of lysine to homoarginine." *Anal Biochem* **287**(1): 110-7.

Hunkapiller, M. W. and L. E. Hood (1983). "Protein sequence analysis: automated microsequencing." *Science* **219**(4585): 650-9.

Hunter, T. C., L. Yang, H. Zhu, V. Majidi, E. M. Bradbury and X. Chen (2001). "Peptide mass mapping constrained with stable isotope-tagged peptides for identification of protein mixtures." *Anal Chem* **73**(20): 4891-902.

Johnson, R. E., C. F. Kirchhoff and H. T. Gaud (2002). "Mannitol-sucrose mixtures-versatile formulations for protein lyophilization." *J Pharm Sci* **91**(4): 914-22.

Kanska, U. and J. Boratynski (2002). "Thermal glycation of proteins by D-glucose and D-fructose." *Arch Immunol Ther Exp (Warsz)* **50**(1): 61-6.

Kaplan, H. and A. Taralp (1997). "Nonaqueous chemical modification of lyophilized proteins." *in* Techniques in Protein Chemistry. D. R. Marshak. London, Academic Press. **VIII:** 219-30.

Karas, M., M. Gluckmann and J. Schafer (2000). "Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors." *J Mass Spectrom* **35**(1): 1-12.

Karas, M. and F. Hillenkamp (1988). "Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons." *Anal Chem* **60**: 2299-2301.

Katayama, H., T. Nagasu and Y. Oda (2001). "Improvement of in-gel digestion protocol for peptide mass fingerprinting by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry." *Rapid Commun Mass Spectrom* **15**(16): 1416-21.

Khmelnitsky, Y. L., S. H. Welch, D. S. Clark and J. S. Dordick (1994). "Salts dramatically enhance activity of enzymes suspended in organic solvents." *J Am Chem Soc* **116**(6): 2647-2648.

Kim, J. S. and H. J. Kim (2001). "Matrix-assisted laser desorption/ionization time-of-flight mass spectrometric observation of a peptide triplet induced by thermal cleavage of cystine." *Rapid Commun Mass Spectrom* **15**(23): 2296-300.

Klibanov, A. M. (1989). "Enzymatic catalysis in anhydrous organic solvents." *Trends Biochem Sci* **14**(4): 141-4.

Klibanov, A. M. (1995). "Enzyme memory. What is remembered and why?" *Nature* **374**(6523): 596.

Klibanov, A. M. (1997). "Why are enzymes less active in organic solvents than in water?" *Trends Biotechnol* **15**(3): 97-101.

Knapp, D. R. (1990). "Chemical derivatization for mass spectrometry." *in* Methods Enzymol. **193:** 314-29.

Knochenmuss, R., A. Stortelder, K. Breuker and R. Zenobi (2000). "Secondary ion-molecule reactions in matrix-assisted laser desorption/ionization." *J Mass Spectrom* **35**(11): 1237-45.

Koskinen, A. M. P. and A. M. Klibanov, Eds. (1996). Enzymatic reactions in organic media. London, Blackie academic & Professional : Chapman & Hall.

Kratzer, R., C. Eckerskorn, M. Karas and F. Lottspeich (1998). "Suppression effects in enzymatic peptide ladder sequencing using ultraviolet - matrix assisted laser desorption/ionization - mass spectormetry." *Electrophoresis* **19**(11): 1910-9.

Krause, E., H. Wenschuh and P. R. Jungblut (1999). "The dominance of arginine-containing peptides in MALDI-derived tryptic mass fingerprints of proteins." *Anal Chem* **71**(19): 4160-5.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R.

Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, J. Szustakowki, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi and Y. J. Chen (2001). "Initial sequencing and analysis of the human genome." *Nature* **409**(6822): 860-921.

Landry, F., C. R. Lombardo and J. W. Smith (2000). "A method for application of samples to matrix-assited Laser desorption ionization time-of-flight targets that enhances peptide detection." *Anal Biochem* **279**: 1-8.

Li, S., T. W. Patapoff, D. Overcashier, C. Hsu, T. H. Nguyen and R. T. Borchardt (1996). "Effects of reducing sugars on the chemical stability of human relaxin in the lyophilized state." *J Pharm Sci* **85**(8): 873-7.

Lundblad, R. L. (1995). Techniques in protein modification. Boca Raton, Fla., CRC Press.

Maillard, L. C. (1912). "Action des acides amines sur les sucres: formation des melanoidins par voie methodologique." *C R Acad Sci* **154**: 66-68.

Mann, M., R. C. Hendrickson and A. Pandey (2001). "Analysis of proteins and proteomes by mass spectrometry." *Annu Rev Biochem* **70**: 437-73.

Matsuo, T. and Y. Seyama (2000). "Special feature: Historical - Introduction to modem biological mass spectrometry (Reprinted from Biological Mass Spectrometry: Present and future)." *J Mass Spectrom* **35**(2): 114-130.

Mazzobre, M. F., M. D. Buera and J. Chirife (1997). "Glass transition and thermal stability of lactase in low-moisture amorphous polymeric matrices." *Biotechnol Prog* **13**(2): 195-199.

Means, G. E. and R. E. Feeney (1971). Chemical modification of proteins. San Francisco,, Holden-Day.

Mishra, P., K. Griebenow and A. M. Klibanov (1996). "Structural basis for the molecular memory of imprinted proteins in anhydrous media." *Biotechnol Bioeng* **52**(5): 609-614.

Monnier, V. M., M. Glomb, A. Elgawish and D. R. Sell (1996). "The mechanism of collagen cross-linking in diabetes: a puzzle nearing resolution." *Diabetes* **45 Suppl 3**: S67-72.

Munch, G., D. Schicktanz, A. Behme, M. Gerlach, P. Riederer, D. Palm and R. Schinzel (1999). "Amino acid specificity of glycation and protein-AGE crosslinking reactivities determined with a dipeptide SPOT library." *Nat Biotechnol* **17**(10): 1006-10.

Munch, G., J. Thome, P. Foley, R. Schinzel and P. Riederer (1997). "Advanced glycation endproducts in ageing and Alzheimer's disease." *Brain Res Brain Res Rev* **23**(1-2): 134-43.

Münchbach, M., M. Quadroni, G. Miotto and P. James (2000). "Quantitation and facilitated de novo sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation-directing moiety." *Anal Chem* **72**(17): 4047-57.

Nakamura, Y., Y. Horii, T. Nishino, H. Shiiki, Y. Sakaguchi, T. Kagoshima, K. Dohi, Z. Makita, H. Vlassara and R. Bucala (1993). "Immunohistochemical localization of advanced glycosylation end products in coronary atheroma and cardiac tissue in diabetes mellitus." *Am J Pathol* **143**(6): 1649-56.

Neglia, C. I., H. J. Cohen, A. R. Garber, P. D. Ellis, S. R. Thorpe and J. W. Baynes (1983). "13C NMR investigation of nonenzymatic glucosylation of protein. Model studies using RNase A." *J Biol Chem* **258**(23): 14279-83.

Neglia, C. I., H. J. Cohen, A. R. Garber, S. R. Thorpe and J. W. Baynes (1985). "Characterization of glycated proteins by 13C NMR spectroscopy. Identification of specific sites of protein modification by glucose." *J Biol Chem* **260**(9): 5406-10.

O'Neil, M. J. and Merck & Co. (2001). The Merck index : an encyclopedia of chemicals, drugs, and biologicals. Whitehouse Station, NJ, Merck.

Opdenakker, G., P. M. Rudd, C. P. Ponting and R. A. Dwek (1993). "Concepts and principles of glycobiology." *Faseb J* **7**(14): 1330-7.

Opella, S. J. and M. H. Frey (1979). "Selection of nonprotonated carbon resonances in solid-state nuclear magnetic resonance." *J Am Chem Soc* **101**: 5854-5856.

Opella, S. J., M. H. Frey and T. A. Cross (1979). "Detection of individual carbon resonances in solid proteins." *J Am Chem Soc* **101**: 5856-5857.

Oppenheimer, H. L., B. Labouesse and G. P. Hess (1966). "Implication of an ionizing group in the control of conformation and activity of chymotrypsin." *J Biol Chem* **241**(11): 2720-30.

Paul, R. G. and A. J. Bailey (1996). "Glycation of collagen: the basis of its central role in the late complications of ageing and diabetes." *Int J Biochem Cell Biol* **28**(12): 1297-310.

Peters, E. C., D. M. Horn, D. C. Tully and A. Brock (2001). "A novel multifunctional labeling reagent for enhanced protein characterization with mass spectrometry." *Rapid Commun Mass Spectrom* **15**(24): 2387-92.

Quan, C. P., S. Wu, N. Dasovich, C. Hsu, T. Patapoff and E. Canova-Davis (1999). "Susceptibility of rhDNase I to glycation in the dry-powder state." *Anal Chem* **71**(20): 4445-54.

Randolph, T. W. (1997). "Phase separation of excipients during lyophilization: effects on protein stability." *J Pharm Sci* **86**(11): 1198-203.

Rice, R. H., G. E. Means and W. D. Brown (1977). "Stabilization of bovine trypsin by reductive methylation." *Biochim Biophys Acta* **492**(2): 316-21.

Roepstorff, P. (2000). "MALDI-TOF mass spectrometry in protein chemistry." *Exs* **88**: 81-97.

Roy, R., E. Katzenellenbogen and H. J. Jennings (1984). "Improved procedures for the conjugation of oligosaccharides to protein by reductive amination." *Can J Biochem Cell Biol* **62**(5): 270-5.

Rudd, P. M. and R. A. Dwek (1997). "Glycosylation: heterogeneity and the 3D structure of proteins." *Crit Rev Biochem Mol Biol* **32**(1): 1-100.

Rupley, J. A. and G. Careri (1991). "Protein hydration and function." *Adv Protein Chem* **41**: 37-172.

Russell, A. J. and A. M. Klibanov (1988). "Inhibitor-induced enzyme activation in organic solvents." *J Biol Chem* **263**(24): 11624-6.

Saeki, H. (1997). "Preparaion of neoglycoprotein from carp myofbrillar protein by maillard reaction with glucose: biochemical properties and emulsifying properties." *J Agric Food Chem* **45**: 680-4.

Sanders, J. K. M. and B. K. Hunter (1987). Modern NMR spectroscopy : a guide for chemists. Oxford [Oxfordshire] ; New York, Oxford University Press.

Sato, R., T. Sawabe, H. Kishimura, K. Hayashi and H. Saeki (2000). "Preparation of neoglycoprotein from carp myofibrillar protein and alginate oligosaccharide: improved solubility in low ionic strength medium." *J Agric Food Chem* **48**(1): 17-21.

Schlack, P. and W. Kumpf (1926). "Uber eine neue methode zue ermitt-lung der constitution con peptiden." *Z Physiol Chem* **154**: 125-170.

Schoellmann, G. and E. Shaw (1963). "Direct evidence for the presence of histidine in the active center of chymotrypsin." *Biochemistry* **2**: 252-5.

Sechi, S. and B. T. Chait (2000). "A method to define the carboxyl terminal of proteins." *Anal Chem* **72**(14): 3374-3378.

Seely, J. H. and N. L. Benoiton (1970). "Effect of N-methylation and chain length on kinetic constants of trypsin substrates. Epsilon-N-methyllysine and homolysine derivatives as substrates." *Can J Biochem* **48**(10): 1122-31.

Shevchenko, A., I. Chernushevich, W. Ens, K. G. Standing, B. Thomson, M. Wilm and M. Mann (1997). "Rapid 'de novo' peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/time-of-flight mass spectrometer." *Rapid Commun Mass Spectrom* **11**(9): 1015-24.

Souillac, P. O., H. R. Costantino, C. R. Middaugh and J. H. Rytting (2002). "Investigation of protein/carbohydrate interactions in the dried state. 1. Calorimetric studies." *J Pharm Sci* **91**(1): 206-16.

Staahl, M., U. Jeppsson-Wistrand, M. O. Maansson and K. Mosbach (1991). "Induced stereo- and substrate selectivity of bioimprinted alpha-chymotrypsin in anhydrous organic media." *J Am Chem Soc* **113**: 9366-9368.

Steitz, T. A. and R. G. Shulman (1982). "Crystallographic and NMR studies of the serine proteases." *Annu Rev Biophys Bioeng* **11**: 419-44.

Stewart, II, T. Thomson and D. Figeys (2001). "18O labeling: a tool for proteomics." *Rapid Commun Mass Spectrom* **15**(24): 2456-65.

Taralp, A. and H. Kaplan (1997). "Chemical modification of lyophilized proteins in nonaqueous environments." *J Protein Chem* **16**(3): 183-93.

Tarelli, E., P. H. Corran, B. R. Bingham, H. Mollison and R. Wait (1994). "Lysine vasopressin undergoes rapid glycation in the presence of reducing sugars." *J Pharm Biomed Anal* **12**(11): 1355-61.

Thiede, B., B. Wittmann-Liebold, M. Bienert and E. Krause (1995). "MALDI-MS for C-terminal sequence determination of peptides and proteins degraded by carboxypeptidase Y and P." *FEBS Lett* **357**(1): 65-9.

Thome, J., J. Kornhuber, G. Munch, R. Schinzel, Y. Taneli, B. Zielke, M. Rosler and P. Riederer (1996). "[New hypothesis on etiopathogenesis of Alzheimer syndrome. Advanced glycation end products (AGEs)]." *Nervenarzt* **67**(11): 924-9.

Triantafyllou, A. O., E. Wehtje, P. Adlercreutz and B. Mattiasson (1997). "How do additives affect enzyme activity and stability in nonaqueous media." *Biotechnol Bioeng* **54**(1): 67-76.

Tsugita, A. (1987). "Developments in protein microsequencing." *Adv Biophys* **23**: 81-113.

Tsukada, H. and D. M. Blow (1985). "Structure of alpha-chymotrypsin refined at 1.68 A resolution." *J Mol Biol* **184**(4): 703-11.

Vakos, H. T. (2001). Applications of the in vacuo chemical modification technique to the study of protein structure and function. Department of Chemistry. Ottawa, University of Ottawa: 409.

Vakos, H. T., B. Black, B. Dawson, M. A. Hefford and H. Kaplan (2001). "In vacuo esterification of carboxyl groups in lyophilized proteins." *J Protein Chem* **20**(6): 521-31.

Vakos, H. T., H. Kaplan, B. Black, B. Dawson and M. A. Hefford (2000). "Use of the pH memory effect in lyophilized proteins to achieve preferential methylation of alpha-amino groups." *J Protein Chem* **19**(3): 231-7.

Vath, J. E., M. Zollinger and K. Biemann (1988). "Method for the derivatization of organic compounds at the sub-nanomole level with reagent vapor." *Fresenius Z Anal Chem* **331**: 248-52.

Vlassara, H. (1996). "Advanced glycation end-products and atherosclerosis." *Ann Med* **28**(5): 419-26.

Vu-Duc, T. and A. Vernay (1992). "A safer methylation procedure with boron trifluoride-methanol reagent for gas chromatographic analysis of ritalinic acid in urine." *J Pharm Biomed Anal* **10**(2-3): 187-91.

Wangikar, P. P., D. Carmichael, D. S. Clark and J. S. Dordick (1995). "Active Site Titration of Serine Proteases in Organic Solvents." *Biotechnol Bioeng* **50**: 329-35.

Watala, C., J. Golanski, H. Witas, R. Gurbiel, K. Gwozdzinski and Z. Trojanowski (1996). "The effects of in vivo and in vitro non-enzymatic glycosylation and glycoxidation on physico-chemical properties of haemoglobin in control and diabetic patients." *Int J Biochem Cell Biol* **28**(12): 1393-403.

Watkins, N. G., S. R. Thorpe and J. W. Baynes (1985). "Glycation of amino groups in protein. Studies on the specificity of modification of RNase by glucose." *J Biol Chem* **260**(19): 10629-36.

Wescott, C. R. and A. M. Klibanov (1994). "The solvent dependence of enzyme specificity." *Biochim Biophys Acta* **1206**(1): 1-9.

Winter, G., A. R. Fersht, A. J. Wilkinson, M. Zoller and M. Smith (1982). "Redesigning enzyme structure by site-directed mutagenesis: tyrosyl tRNA synthetase and ATP binding." *Nature* **299**(5885): 756-8.

Wirth, D. D., S. W. Baertschi, R. A. Johnson, S. R. Maple, M. S. Miller, D. K. Hallenbeck and S. M. Gregg (1998). "Maillard reaction of lactose and fluxetine hydrochloride, a secondary amine." *J Pharm Sci* **87**: 31-39.

www.ncbi.nlm.nih.gov.

Xu, K. and A. M. Klibanov (1996). "Ph Control of the Catalytic Activity of Cross-Linked Enzyme Crystals in Organic Solvents." *J Am Chem Soc* **118**(41): 9815-9819.

Yaylayan, V. A. and A. Huyghues-Despointes (1994). "Chemistry of Amadori rearrangement products: analysis, synthesis, kinetics, reactions, and spectroscopic properties." *Crit Rev Food Sci Nutr* **34**(4): 321-69.

Yeboah, F. K., I. Alli and V. A. Yaylayan (1999). "Reactivities of D-glucose and D-fructose during glycation of bovine serum albumin." *J Agric Food Chem* **47**(8): 3164-72.

Yennawar, H. P., N. H. Yennawar and G. K. Farber (1995). "A structural explanation for enzyme memory in nonaqueous solvents." *J Am Chem Soc* **117**(2): 577-585.

Zacharis, E., P. J. Halling and D. G. Rees (1999). "Volatile buffers can override the "pH memory" of subtilisin catalysis in organic media." *Proc Natl Acad Sci U S A* **96**(4): 1201-5.

Zaks, A. and A. M. Klibanov (1988). "The effect of water on enzyme action in organic media." *J Biol Chem* **263**(17): 8017-21.

Zaks, A. and A. M. Klibanov (1988). "Enzymatic catalysis in nonaqueous solvents." *J Biol Chem* **263**(7): 3194-201.

Zoller, M. J. and M. Smith (1982). "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA." *Nucleic Acids Res* **10**(20): 6487-500.

Zoller, M. J. and M. Smith (1983). "Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors." *Methods Enzymol* **100**: 468-500.

Publications, Patents and Conference Proceedings

Publications:

Stewart, N. A., A. Taralp and H. Kaplan (1997). "Imprinting of lyophilized alpha-chymotrypsin affects the reactivity of the active-site imidazole." *Biochem Biophys Res Commun* **240**(1): 27-31.

Stewart, N. A., V. T. Pham, C. T. Choma and H. Kaplan (2002). "Improved peptide detection with matrix-assisted laser desorption/ionization mass spectrometry by trimethylation of amino groups." *Rapid Commun Mass Spectrom* **16**(15): 1448-53.

Patents:

"Increased sensitivity of peptide detection with Matrix-Assisted Laser Desorption/Ionization Mass spectrometry by in vacuo methylation of amino groups", co-inventor with H. Kaplan., Canadian Patent No. CA 2377446, US Patent Application Publication No. US 2003/0180962 A1

"In Vacuo glycation of proteins", co-inventor with M. King and H. Kaplan., Canadian Patent No.CA 2398213.

Conference Proceedings:

Protein Society 16th Symposium, August 17-21, 2002, San Diego, CA.
"*In Vacuo* glycation as an approach to covalent cross-linking of proteins" M.C. King, H. Kaplan, N.A.S. Stewart.

Protein Society 14th Symposium, August 5-9, 2000, San Diego, CA.
"Selective isolation of peptides from enzymatic digests as an approach for sequencing the N and C-terminal ends of proteins" H. Kaplan, N.A.S. Stewart, V.T. Pham, T. Cyr and M.A. Hefford.

Ottawa Carleton Chemistry Institute (OCCI) 2000
"Non-aqueous glycosylation of lyophilized proteins" N.A.S. Stewart, Paul Baba, and H. Kaplan.

Protein Society 11th Symposium, July 12-16, 1997, Boston, MA.
"Ligand-induced changes in the active site of lyophilized chymotrypsin" N.A.S. Stewart and H. Kaplan.

OCCI 1997
"Chemical modification of lyophilized α-chymotrypsin: protection studies" N.A.S. Stewart, A. Taralp and H. Kaplan.

Claims to Original Research

1. Iodomethane was reacted with lyophilized peptides and proteins *in vacuo* to produce a trimethylammonium derivative of their amino groups for a significant increase in MALDI detection.
2. *In vacuo* methylation of a lyophilized enzyme, α-chymotrypsin, in the presence and absence of imprinting ligands to probe the enzyme's active site.
3. Preserving the activity of an enzyme from inactivation by iodomethane by the co-lyophilization with imprinting ligands.
4. The application of the *in vacuo* methylation procedure for the isolation of C-terminal peptides derived from an enzymatic digest of a protein.
5. A miniature high voltage paper electrophoresis apparatus was designed and successfully employed to separate C-terminal peptides from enzymatic digest of a protein.
6. Extensive glycation of proteins with a reducing sugar by heating a co-lyophilized mixture of a protein and glucose *in vacuo*.
7. The predominant formation of the cyclic ketoamine as the product from the derivatization of amino groups of proteins by *in vacuo* glycation.