This is to certify that the
dissertation entitled

STUDY OF PRIMARY AMINES FOR OPTIMUM NUCLEOPHILIC CLEAVAGE OF CYANYLATED CYSTEINYL PROTEINS DURING THE DEVELOPMENT OF AN 'ON-LINE' SYSTEM FOR THE DISULFIDE MASS MAPPING OF CYSTINYL PROTEINS

presented by

JOSE LUIS GALLEGOS-PEREZ

has been accepted towards fulfillment
of the requirements for the

PhD degree in CHEMISTRY

Major Professor's Signature

14 December 2005

Date

STUDY OF PRIMARY AMINES FOR OPTIMUM NUCLEOPHILIC CLEAVAGE OF CYANYLATED CYSTEINYL PROTEINS DURING THE DEVELOPMENT OF AN 'ON-LINE' SYSTEM FOR THE DISULFIDE MASS MAPPING OF CYSTINYL PROTEINS

By

José Luis Gallegos-Pérez

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

2005

ABSTRACT

STUDY OF PRIMARY AMINES FOR OPTIMUM NUCLEOPHILIC CLEAVAGE OF CYANYLATED CYSTEINYL PROTEINS DURING THE DEVELOPMENT OF AN 'ON-LINE' SYSTEM FOR THE DISULFIDE MASS MAPPING OF CYSTINYL PROTEINS

By

José Luis Gallegos Pérez

Undoubtedly, modern mass spectrometry combined with chemical derivatization is a very powerful tool in the analysis of biomolecules; currently, it is the most important analytical tool used in proteomics. In this dissertation, the use of mass spectrometry in the cyanylation-based disulfide mass mapping gives us an excellent example of the application of mass spectrometry in biomedical science and in the determination of the disulfide proteome. The cyanylation-based disulfide mass mapping methodology relies on the chemical derivatization and cleavage of the protein prior to analysis by mass spectrometry.

In this dissertation, the results of a quantitative and qualitative study of the nucleophile-mediated CN-induced cleavage reaction is presented in the context of the cyanylation-based methodology. The focus is on the use of primary (methyl to butyl) amines as an alternative to ammonia as the cleavage reagent. During optimization of chemical cleavage of cyanylated cysteinyl proteins and the detection of the resulting cleavage products, several parameters such as type of nucleophile, concentration, temperature, and reaction time were systematically

studied using Ribonuclease A as a model protein. High performance liquid chromatography was used to monitor the yield of the expected cleavage products as identified by mass spectrometry.

Also presented is the use of binary mixtures of homologous nucleophiles to improve the confidence in distinguishing CN-induced cleavage products from chemical by-products during analyses of cleavage reaction mixtures by MALDI-MS. The use of two homologous nucleophiles for the cleavage reaction provomotes the appearance of pairs of mass spectral peaks that facilitate recognition of diagnostic CN-induced cleavage products, and increases confidence in results when samples present weak signals or interfering chemical noise. The analytical advantage that can be gained in the use of a binary mixture of homologous nucleophiles for cyanylation-induced cleavage of partially reduced isoforms of cystinyl proteins is evaluated, and the combined use of ammonia and methylamine is proposed for labeling of cleavage products in a single chemical step.

Finally, a preliminary development of an 'on-line' system for the cyanylation-based methodology is completely evaluated as a mean of carrying out all the chemistry 'on-line' to avoid sample manipulation and, by consequence, possible sample losses associated with stepwise batch processing.

DEDICATION

A Dios y la Virgen de Guadalupe a quien me encomendé al inicio de esta etapa en mi vida.

(To God and Our Lady of Gualaupe whom I entrusted when I started this stage in my life.)

A mi madre María Gallegos Pérez. Gracias por tus enseñanzas y amor.

(To my mother María Gallegos Pérez. Thank you for your teachings and your love.)

A quien sigo esperando llegue a mi vida.

(To the one for whom I still wait.)

ACKNOWLEDGMENTS

I want to express my gratitude to Michigan State University (MSU) not only for the excellent academic graduate program in which I was involved but also for the human values demonstrated.

My gratitude to the community founded by the Saint Marcelino Champagnat, the Marist Brothers. Thank you also to Querétaro Autonomous University and the México National Autonomous University.

The fulfillment of this stage in my life would not have been possible without the advice and support of Dr. Jack Throck Watson, my advisor. Thank you very much for your understanding when I had some hard times in my life. Thank you for the time you spent to teach me or just to listen when I needed it. Special mention to Mrs. Shirley Watson, thanks for helping me with my English writing.

Thank you to my committee members: Dr. Merlin Bruening, Dr. Marcus Dantus and Dr. Greg Baker.

I also want to thank to people that always will remain in my heart: Tita Carmen, Tita Amalia, Sandra Verónica, Pbro. F. Carrillo, Sra. Carmen Ortega and her husband, grandparents, godparents, aunts, cousins and the rest of my family, especially Madrina Elena, Tía Remedios, Mago, Perla, Chava and wife.

I want to thank all my friends in México and other countries. Thank you for all your support and encouragement when I needed it. Thanks especially to the families de la Guardia Herrera, Aguado Hernández and Ibarra Montaño for your help before and through all my stay at MSU.

I also thank my friends that I met at MSU: Eliot, Sandra, and Sarhi; Kai-Chun, Vitor and Sandra, Gabe, Eugenia, Tannia, Kiran, Paolo, Yongjae, Arindom, Devi and family, Arun and family, Cchimee and Sanjeev.

Thanks to previous and current Watson group members: Theresa Burks, Stephen Robert Bowman, Jianfeng (Frank) Qi, Wei Wu, Chad Borges, Yingda Xu, Laura Rangel Ordóñez, Charles Ngowe, Dr. Robin J. Hood, Yi-Te (Roberto) Chou, Xue Li (now Dr. Shelly), Jamie Dunn and Olivia Guerre; to mass spec facility personnel, Susi, Rhonda, Mrs. Bev, Lijun for your friendship and support

To anyone else not mentioned thank you.

Last but not least to Laura, you have shared as a friend all my stay at MSU since I applied to come to USA. I have cherished all your support, first from México, then during your stay at MSU, and now even from Germany during good and bad times.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AC	Alternating Current
APCI	Atmospheric Pressure Chemical Ionization
BDK	Bradykinin
CAD	Collisionally Activated Dissociation
CDAP	1-cyano-4-dimethylamino-pyridinium
CID	Collision-Induced Dissociation
CN	Cyano group
CNBr	Cyanogen Bromide
Cys	Cysteine residue
Da	Dalton
DC	Direct Current
ECD	Electron Capture Dissociation
ESI	Electrospray Ionization
FAB	Fast Atom Bombardment
FSA	Fragment Sets Algorithm
FT	Fourier Transform
HPLC	High Performance Liquid Chromatography
ICAT	Isotope-Coded Affinity Tag
ICR	Ion Cyclotron Resonance
IF1…IF4	Cyanylated single-reduced isoforms of Ribonuclease A (numbered)
IF1L…IF4L	Cyanylated single-reduced isoforms of α-Lactalbumin (numbered)
itz	Iminothiazolidine-4-carboxyl terminus
LIT	Linear Ion Trap
m/z	Mass-to-Charge ratio
MALDI	Matrix Assisted Laser Desorption/Ionization
MCAT	Mass-Coded Abundace Tagging
$MeNH_2$	Methylamine

MH^+	Protonated molecule
MRM	Multiple Reaction Monitoring
MS	Mass Spectrometry
MS/MS	Mass Spectrometry /Mass spectrometry (tandem mass spectrometry)
MS^n	When used to describe the mass spectrum that results from the controlled dissociation of a precursor ion of a series of 1^{st}, 2^{nd}, etc. generation product ions (various mass spectrometry/mas spectrometry techniques) (from reference 39 in Chapter 1)
NSA	Negative Signature Algorithm
oa-TOF	Orthogonal Acceleration Time Of Flight
PD	Plasma Desorption
PRCC	Partial Reduction, Cyanylation and Cleavage (Methodology for Disulfide Mass Mapping)
Q	Quadrupole Mass Analyzer
q	Quadrupole used as collision cell no for mass separation
QIT	Quadrupole Ion Trap
QUISTOR	Quadrupole Ion Storage
RNase-A	Ribonuclease-A
S/N	Signal-to-Noise ratio
SIM	Selective Ion Monitoring
TCEP	Tris(2-carboxyethyl) phosphine
TFA	Trifluoroacetic acid
TOF	Time Of Flight
α-CHCA	α-cyano-4-hydroxy cinnamic acid
β	β-elimination of HSCN from cyanylated cysteine residue
μ-HPLC	Micro-High Performance Liquid Chromatography

CHAPTER 1 DISULFIDE MASS MAPPING OF PROTEINS

1.- PROTEINS

1.1 Biological importance of proteins

Proteins are essential constituents of all organisms. Some proteins can store and transport a variety of particles ranging from macromolecules to electrons. In some cases, membrane proteins guide the flow of electrons in the vital process of photosynthesis and transmit information between specific cells and organs in complex organisms. Virtually every property that characterizes a living organism is affected by proteins [1, 2]. Some proteins provide the structure that helps give cells their bioactive shape; others serve as hormones to carry signals from one cell to another. Another protein, hemoglobin, carries oxygen from the lungs to the rest of the body. And perhaps most important of all, proteins serve as enzymes that catalyze thousands of chemical reactions necessary to sustain life. The diversity of protein structures and functions is as wide as the diversity of the life itself [3].

The structure-function paradigm states that a specific function of a protein is determined by its unique and ordered three-dimensional structure; this is one of the cornerstones in protein biology and biochemistry (see Figure 1). These days, scientists believe in the existence of "disordered" proteins, which in their

purified state at neutral pH lack an ordered structure [4, 5]. In summary, the knowledge of the protein structure is an important issue in the protein study field [5].

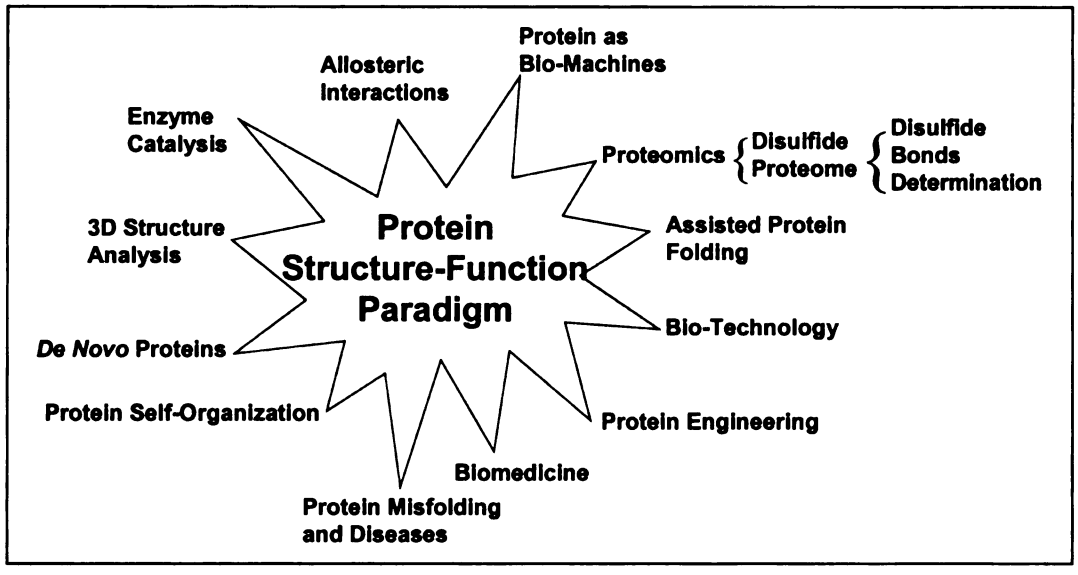


Figure 1. Protein-Structure Paradigm and the Disulfide Proteome (from [5]).

2.- PROTEOMICS AND THE DISULFIDE PROTEOME

2.1 Introduction

All the information necessary to make a complete human being is inside each of our cells. However, not all genes are expressed in all cells. All cells express some genes whose protein products provide unique cell-specific functions. A "proteome" is the entire PROTEin complement expressed by a genOME, or by a cell or tissue type; it represents the collection of proteins under specific conditions. As an extrapolation of the concept of the "genome project", a

"proteome project" is research that seeks to identify and characterize the proteins present in a cell or tissue and define their patterns of expression [6, 7]. In other words, proteomics is the study of multiprotein systems, in which the focus is on the interplay of multiple, distinct proteins in their roles as part of a larger system such as a cell line, tissue, or complete organism [6, 8]. It is said that protein chemistry and proteomics involve protein identification, but the substantial difference between them is that the goal of proteomics is to characterize the behavior of a system rather than the behavior of a single component such as in protein chemistry.

Science is moving from genomics to proteomics in order to get insight into the functional network of gene expression. Actually, proteomics is older than genomics. The first studies that can be called proteomics began in 1975 with the introduction of a two-dimensional gel by several research groups who began mapping proteins from *Escherichia coli*, mice, and guinea pigs. At that time, around 300 proteins could be separated, but due to technical limitations, they could not be completely identified. Currently, with more refined techniques, up to 10,000 protein spots have been separated under highly reproducible conditions. [9, 10].

Recently, the concept of systems biology has been introduced in the study of living systems from a holistic perspective based on the profiling of a multitude of biochemical components (see Figure 2). From this point of view, proteomics and

genomics are a part of a whole in which the final goal is the study of a living system.

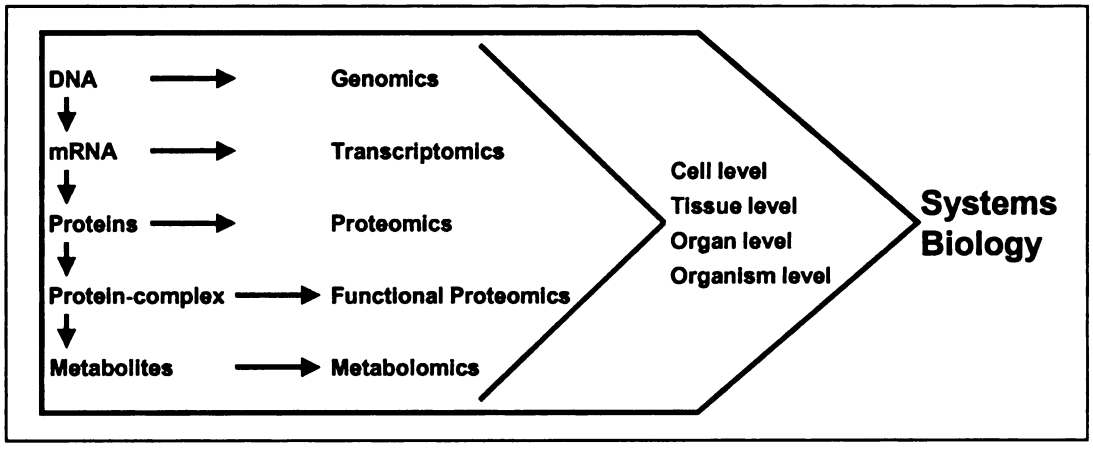


Figure 2. Different levels of measurement in a system approach. (Adapted from [11])

2.2 Proteomics and Biology

The accumulation of complete sequences of genomes is not sufficient to elucidate biological function. From the study of genes solely, not all types of information can be obtained. For example, proteins are responsible for the phenotypes of cells, not genes. It is impossible to elucidate mechanisms about diseases, aging and environmental effects solely by studying the genome. Thus, the importance of proteomics is a direct result of advances made in large-scale nucleotide sequencing of expressed sequence tags and genomic DNA. It is only through the study of the proteins that questions about biological functions can be answered [9, 12].

Four principal applications constitute protemics:

1) Mining: This is the simple identification of as many of the proteins in a sample as possible.

2) Protein-expression profiling: This is the identification of a protein in a sample in a particular state of the organism or cell, e. g., in a disease state or as a function of exposure to a drug or other stimulus. Normal and diseased cells can be compared to determine which proteins are expressed differently in different states.

3) Protein-network mapping: This is the proteomics approach to determining how proteins interact with each other in living systems.

4) Mapping of protein modifications: In this part of proteomics, the task consists of identifying how and where proteins are modified. Postranslational modifications usually govern the structure and function of proteins, which is why it is important to know how a protein has been modified.

In recent years, attention has shifted to "functional proteomics" where the proteins are purified with a specific function in mind. An example of this is the purification of an organelle or of a multi-protein complex with a particular function [8].

2.3 The Disulfide Proteome and biological importance of disulfide bonds

The disulfide proteome is the subset of proteins containing disulfide bonds in a proteome. The thiol redox status adds a new dimension to biology as well as medicine and any field related with protein science [13].

In 1961, the work of Anfinsen et al., showed that disulfide bonds are formed spontaneously, in RNase-A, upon oxidation in air. This work showed that a polypeptidic chain has to be folded before it becomes a biologically active protein. This experiment was the first to show that a protein could fold into its correct tertiary structure with only the information provided in its aminoacid sequence [14]. Since then, the two major facts, which are central to understanding the mechanism of protein folding are being studied. They are: 1) the development of the theory of folding, especially for predicting structure from sequence, and 2) analysis of the relation between sequence homology and evolutionary relatedness [15, 16].

Disulfide bonds are one of the key aspects related to folding among cystinyl proteins and, consequently, they are important for the understanding of the chemical structure and function of such a protein. Disulfide bonds are important for the stabilization of the native structures of cystinyl proteins, emerging as the primary mechanism by which the activity of a protein is

established [13, 17]. Without disulfide bonds, cystinyl proteins may be disordered, weak and non-functional [18]. Two classes of disulfide-bonded proteins are known: the first, in which the inter-cysteine linkage is a stable part of the final folded structure, and the second, in which cysteine pairs alternate between the reduced and the oxidized states.

In organisms from *Escherichia coli* to humans, disulfide bond formation is important to proper folding of secreted protein, such as bacterial virulence factors and mammalian glycoproteins [19]. In toxicology, the knowledge of cysteine linkage is of great importance especially for biologically active peptides derived from venomous animals and toxin-producing bacteria, where peptide toxins can possess multiple disulfide bonds. One good example of toxins with disulfide bonds is the neurotoxins derived from marine snails of the genus *Conus*, which belong to a family of bioactive peptides called conotoxins or conopeptides. These short, disulfide-rich peptide neurotoxins are produced in the venom of predatory marine cone snails. It is generally accepted that an estimated 100,000 unique conotoxins fall into only a handful of structural groups, based on their disulfide bridging frameworks. Conopeptides have defined tertiary structures formed and stabilized by their disulfide bonds [20, 21]. Toxins from spiders, snakes, or scorpions also show polypeptidic structures containing disulfide bonds [22-24].

In some diseases, such as the prion transmissible spongiform encephalopathies, the key event seems to be the conversion of the prion protein (PsP) from its normal cellular form (PrP^c) to an abnormal scrapie isoform (PrP^{sc}) that may involve a covalent reaction of the sole intramolecular disulfide bond of PrP^c [25]. Involvement of disulfide bonds in stabilization of structure and function of the Newcastle disease virus hemagglutinin-neuraminidase has also been published [26, 27].

The disulfide proteome is also gaining importance because of the evidence suggesting that redox regulation plays a critical role in cell biology. The redox cycling of the disulfide bond, its formation and dissociation, may be central to an enzyme's activity or to protein activation allowing for a "biological switch" that regulates the protein activity [28-30]. For example, in plants, reduction of regulatory disulfide bonds using reducing power generated by light, activates the regulatory photosynthetic enzymes and deactivates glucose-6-phosphate dehydrogenase (G6PDH) –an enzyme involved in the oxidative pentose phosphate pathway [31].

Therefore, we can conclude that in the genomic era, when the proteomes of hundreds of different organisms are known, it is particularly interesting to be able to highlight the "disulfide proteome" (see Figure 3) for the comprehensive analysis of the cell function and regulation [19].

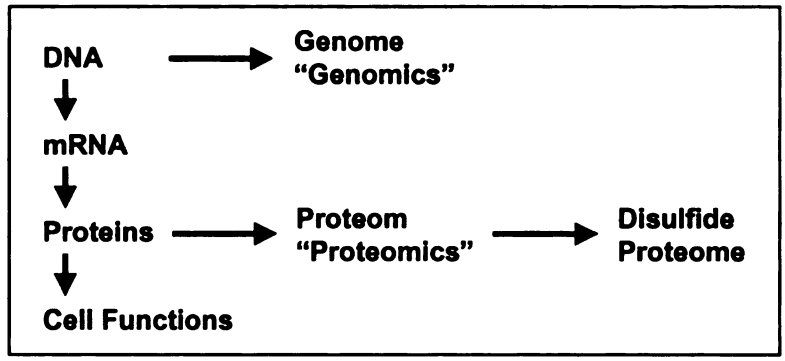


Figure 3. Biochemical context of genomics, proteomics and the disulfide proteome (adapted from [6]).

3.- MASS SPECTROMETRY AND PROTEIN CHARACTERIZATION

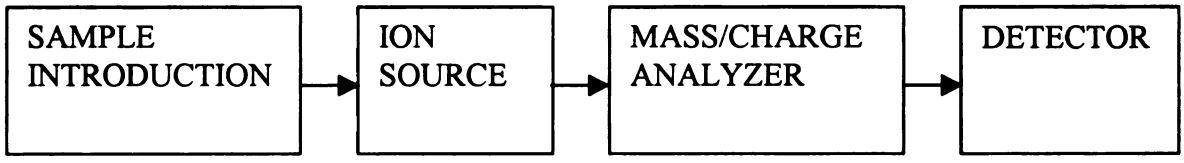
3.1 Introduction

During the last two decades of the last century, the need for structural and quantitative information from biological samples led to the development of ionization methods capable of handling fragile, nonvolatile, high-mass molecules, avoiding, if possible, any derivatization step [32]. Currently, the determination of the mass of biological compounds is one of the most important pieces of experimental data in the analysis of their structure.

There is not doubt that mass spectrometry is the most accurate method to determine the molecular mass of biological molecules. For a long time, chemists and biochemists were searching for a simple method to determine accurately and sensitively the molecular masses of biopolymers such as proteins or polysaccharides [33]. The application of mass spectrometry to the study of

biology was realized in the early 1980s with the advent of reliable ionization methods such as Fast Atom Bombardment (FAB) and Plasma Desorption (PD). This made possible the ionization of small peptides and proteins. However, the two developments in mass spectrometry (MS) instrumentation that revolutionized protein chemistry were the introduction of electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) [34, 35]. These methods solved the problem of generating ions from biomolecules. Recently, Electron Capture Dissociation (ECD) has emerged as a new technique that can be applied to peptide and protein sequencing.

All mass spectrometers have four essential components:



The sample introduction, depending on the ionization technique, could be gas chromatography, liquid chromatography, electrophoresis (e.g capillary electrophoresis, isotacophoresis, and other related techniques) or a direct insertion probe. Molecules of the sample are converted to ions in the ion source, and the produced ions are separated according to their mass/charge ratio (*m/z*). The ion beam coming from the mass analyzer is detected and transformed into a usable signal by a detector.

Currently, many different ionization techniques and mass analyzers are used in mass spectrometry, but in general the instrumentation layout is the same [36].

Table 1. Overview of the selected ionization techniques used in mass spectrometry. (Adapted from reference [33] with additions from [37-40]).

METHOD	GENERAL	ADVANTAGES	DRAWBACKS	MASS LIMIT
Gas-phase ionization				
Electron Ionization (EI)	High-energy electrons (70 eV) interact with molecules in gas phase provoking ionization.	Applicable to nearly all volatile compounds. Easily coupled to separation techniques (such as Gas Chromatography) Reproducible fragmentation pattern.	Compound must be thermally volatile and stable. Molecular ion is sometimes not detectable.	~ 700 Da
Chemical Ionization (CI)	Ionization of the analyte occurs as result on an ion/molecule reaction. A reagent gas (e.g., ammonia, methane) is ionized by an electron beam and resulting ions react with neutral analyte molecule to produce analyte ions.	Protonated molecule is often detectable; little fragmentation observed.	Compound must be thermally stable and volatile. Mass spectra depend on the type of reagent gas used.	~ 700 Da
Particle bombardment				
Fast Atom Bombardment	Sample is dissolved in a liquid matrix and ionized by bombardment with high-energy atom beam.	Rapid, simple. Strong ion current that can be useful for high-resolution measurements.	Analyte must be soluble in the matrix. High chemical background. Not good for analytes with more than 2 charges.	~ 6000 Da
Secondary-Ion mass spectrometry (SIMS)	Ionization occurs through bombardment of a solid sample with a high-energy beam of ions instead of fast atoms as for FAB.	Same as for FAB. Higher sensitivity for higher masses.	Similar to FAB.	~ 13000 Da

METHOD	GENERAL	ADVANTAGES	DRAWBACKS	MASS LIMIT
Laser Desorption				
Matrix-Assisted Laser Desorption/Ionization	A laser is used to produce ions from an analyte that is mixed in a solid matrix. The laser provides energy for the ionization and the desorption of the ions from the matrix.	Tolerates moderate quantities of impurities.	MS/MS is difficult, no easy coupling with LC, requires a mass analyzer that is compatible with pulsed ionization techniques. Low reproducibility.	~ 500 000 Da
Atmospheric Pressure Ionization				
Electrospray (Nanoelectrospray is a low-flow electrospray technique that does not use a make-up solvent where the sample flow is dependent on the potential on the tip of the electrospray needle. It is known commercially as NanoSpray, MicrolonSpray or NanoFlow)	The sample solution is sprayed into a high electric field (a few kilovolts) from a needle.	Easy couple to LC; low background.	Very sensitive to salts. Multiply charged species are formed.	~ 200 000 Da
Electron Capture Dissociation (ECD)	A hybrid "soft" ionization technique where low-energy electrons (< 0.2 eV) are captured by protein cations; the resulting radical-cation dissociates extensively	ECD preserves the secondary structure and the labile groups in a polypeptide chain, which facilitates the assignment of post-translational modifications.	Currently, it has only been coupled to Fourier Transform ion cyclotron resonance mass spectrometers.	Up to 42 000 Da has been reported with improved methodology

3.2 Ionization techniques

3.2a Fast Atom Bombardment (FAB)

During the 1970s, several techniques were introduced to obtain mass spectra of compounds with negligible vapor pressure. However, the technique with the greatest impact on the analysis of nonvolatile and/or thermally labile compounds was FAB. FAB was developed by Barber and co-workers, who published the novel procedure of introducing the analyte in a viscous liquid (matrix). Since that time, FAB has been used to analyze a wide variety of molecules ranging from oligopeptides to organometallics and buckminsterfullerenes. [38, 41]

In FAB, a small amount of sample (matrix + analyte) is placed on a target, which is subsequently bombarded with a fast atom beam (for example, 5-8 keV of neutral xenon or argon atoms). The impact desorbs protonated molecules (ions) and fragments from the analyte (see Figure 4). Cluster ions from the liquid matrix are also desorbed and produce a chemical background that varies with the matrix used. There are several liquid matrixes, the more common of which are glycerol, thioglycerol, m-nitrobenzyl alcohol, and diethanolamine [37]. There is a variation of the technique termed "liquid secondary ion mass spectrometry" (LSIMS). In this variation, cesium ions (Cs^+) are used instead of a neutral atom beam; it seems that the charge of the primary bombarding beam is of little

importance in the production of sample ions. What does seem to be of major importance is the momentum of the bombarding particle. The momentum can be varied by either increasing the acceleration energy of the particle or by increasing the mass of the bombarding particle (Xe > Ar) [41].

The mechanism of FAB is not well understood, but the FAB mass spectra are likely to be dominated by cluster ions from the matrix. Cationized adducts of the sample are observed: $[M+H]^+$, $[M+Na]^+$, $[M+K]^+$ as well as negatively charged ions, $[M-H]^-$. For monovalent ionic salts, in a positive FAB spectrum the cation C^+ is observed as well as cluster ions $[C_n + X_{n-1}]^+$ and in the case of a negative FAB mass spectrum, the presence of ion X^- and clusters $[C_{n-1} + X_n]^-$ are evident.

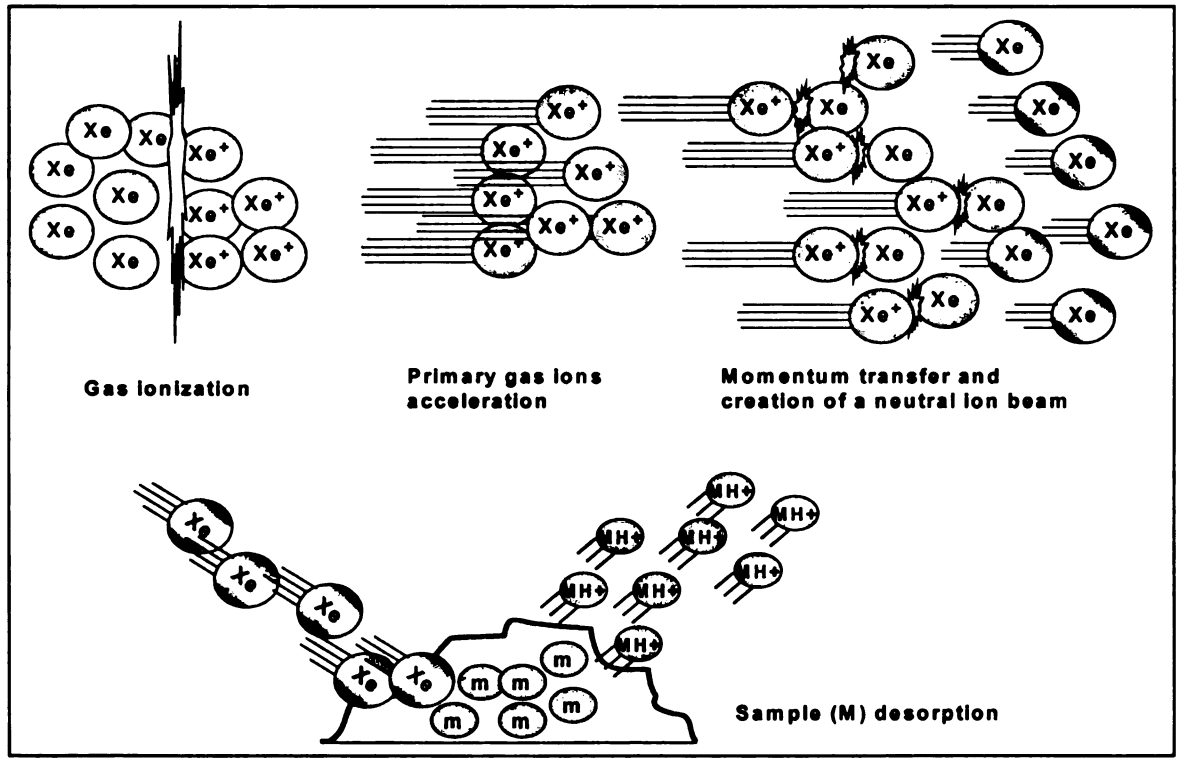


Figure 4. Fast atom bombardment process. In the scheme Xenon gas is used to desorb sample M.

3.2b Matrix-Assisted Laser/Desorption Ionization (MALDI)

MALDI was developed by Hillenkamp and Karas during the late 1980s and constituted a milestone in the analysis of biomolecules [35, 42-44]. Currently, it is nearly impossible to open a recent issue of any journal in the field of protein science or biochemistry without finding at least one paper on MALDI (original paper on application in biomolecules has been cited at least 1369 times according to SciFinder Scholar© accessed July 31, 2005 [35]). A successful MALDI analysis encompasses several crucial steps: sample preparation, excitation of the sample and disintegration of the condensed phase, generation and separation of charges and ionization of analyte molecules, and, finally, extraction and separation according to the *m/z* value of the ions in the mass spectrometer and detection. The technique makes use of the absorption of laser light by a solid sample layer. In MALDI, the analyte is mixed with a 100-50,000-fold molar excess of small, UV-absorbing molecules (the matrix) to form a solid-phase deposit on a sample plate.

MALDI matrices must meet a number of requirements:

- be capable of forming a homogeneous and fine crystalline solid during the sample preparation.
- be able to embed and isolate analytes (e.g., by co-crystallization)
- be soluble in solvents compatible with analyte

- be vacuum stable

- have a high absorptivity for the laser radiation

- cause co-desorption of the analyte upon laser irradiation

- promote analyte ionization [45]

The matrix absorbs and transfers the energy from the laser pulse to the sample producing a plasma that allows in vaporization and ionization of the analyte. The ionization process produces both positive and negative ions. The analyte molecules usually combine with just one proton to form single-charge ions $[M+H]^+$; however, some double- and triple-charge ions may be observed (e.g., $[M+2H]^{2+}$, $[M+3H]^{3+}$). The presence of sodium and potassium adducts is also possible (Figure 5).

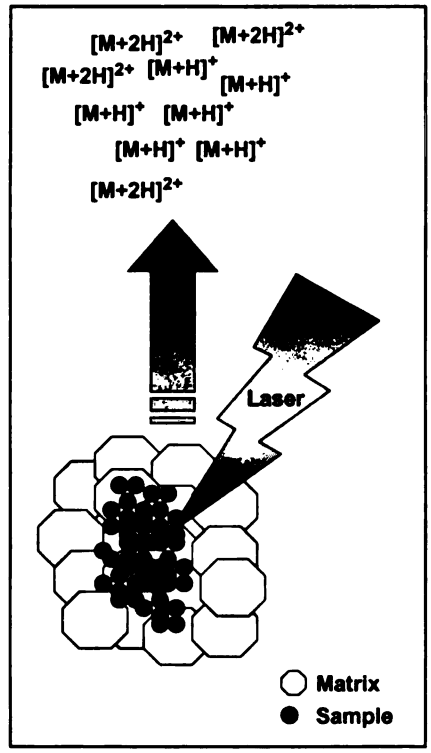


Figure 5. Sketch of the MALDI desorption process.

MALDI is a complex chemical event, where one laser shot/one mass spectrum is obtained in few nanoseconds. After the laser shot, a plume of material containing neutral matrix molecules, as well as reactive species (matrix radicals, electrons, and hydrogen atoms), is formed. Within this plume, regardless of a low average degree of ionization, suitable analytes can be ionized efficiently.

Models for MALDI were developed considering three different steps:

1) Incorporation and isolation of analyte molecules in the host matrix

2) Release of molecules into the vacuum upon disintegration of the matrix-analyte solid after laser-energy deposition (energy absorption)

3) Ionization by ion-molecule reaction in the laser plume including matrix reactive species.

These different steps frequently have been often considered and studied separately isolated, i.e., the transition from the solid state to the gas phase from a physical point of view or the ionization from a chemical or thermodynamic one. One major consequence of the separation in single mechanistic step was that the details of the solid-gas phase transition were not considered with respect to their relevance for ionization. As a consequence, a "photochemical" ionization mechanism was favored in the initial MALDI papers, considering the generation of neutral analyte species as the first step and then ionization by proton-transfer

coming from matrix species in the gas phase matrix plume. Seldom were other possible suggested mechanisms taken into account [46-48].

A major advance in models of MALDI was to include cluster formation and generation of charged clusters by either charge separation (deficit/excess of anions or cations) or photochemical processes as an important primary ionization step. This model (the "lucky survivor" model) tries to be comprehensive and general for MALDI-MS ionization [47]; the key point of this model is to consider ionization during desorption as a dominant ionization and charge-separation pathway. Recently, a quantitative model of ionization in ultraviolet MALDI was presented. This model was focused initially on primary ionization in a MALDI sample containing only matrix. Later, it was extended to include ion-molecule reactions [49, 50].

However, even after all these proposals for MALDI models, the mechanism for ionization in MALDI is still under investigation [38].

From a practical point of view, MALDI sample preparation techniques are vital to forming a homogenous solid sample that yielded imporved laser shot-to-shot reproducibility. Sample preparation and recent improvements are given elsewhere [45, 51]. Development of free-matrix solid surfaces where the sample is directly deposited for analysis have also been developed [52, 53]. Although MALDI is not easily coupled to separation methodologies, this technique offers

several advantages: 1) Biological samples can be examined without extensive purification because MALDI is somewhat tolerant of buffers and salts, (although in order to allow the desorption of proteins and DNA in a large excess of salt or surfactant, different MALDI probes have been recommended [52-56]). 2) It has a low sample requirement, usually 1-10 pmol of analyte is enough to obtain a good quality mass spectra. 3) Proteins with masses ranging to greater than 100 kDa can be analyzed.

3.2c Electrospray Ionization (ESI)

The development of electrospray ionization (ESI) started with the research of Dole et al. in the 1960s who described the idea that electrospraying a liquid containing analyte molecules might liberate these as ions and make them amenable to mass spectrometry [57, 58]. Fenn et al., in 1984 [59-61] demonstrated the potential of ESI as a mass spectrometry technique for small and large biomolecules [34, 62]. Independently and almost at the same time, Aleksandrov et al, coupled a liquid chromatography micro-column to a mass spectrometer where the ions were extracted from the eluent solution by spraying into a non-uniform electric field. They observed protonated and cationized products of small molecules [63, 64].

Electrospray ionization (ESI) is a soft ionization technique in which ions that are in solution are put into the gas phase through a mechanism of ion

desorption or ion evaporation. The sample solution is sprayed across a high potential difference (a few kilovolts, 3-8 kV) from a narrow-bore needle (0.1-0.3 mm) into an orifice in the interface. In conventional ESI, heat and gas (N_2) flows are used to desolvate the ions existing in microscopic droplets of the sample solution (See Figure 6).

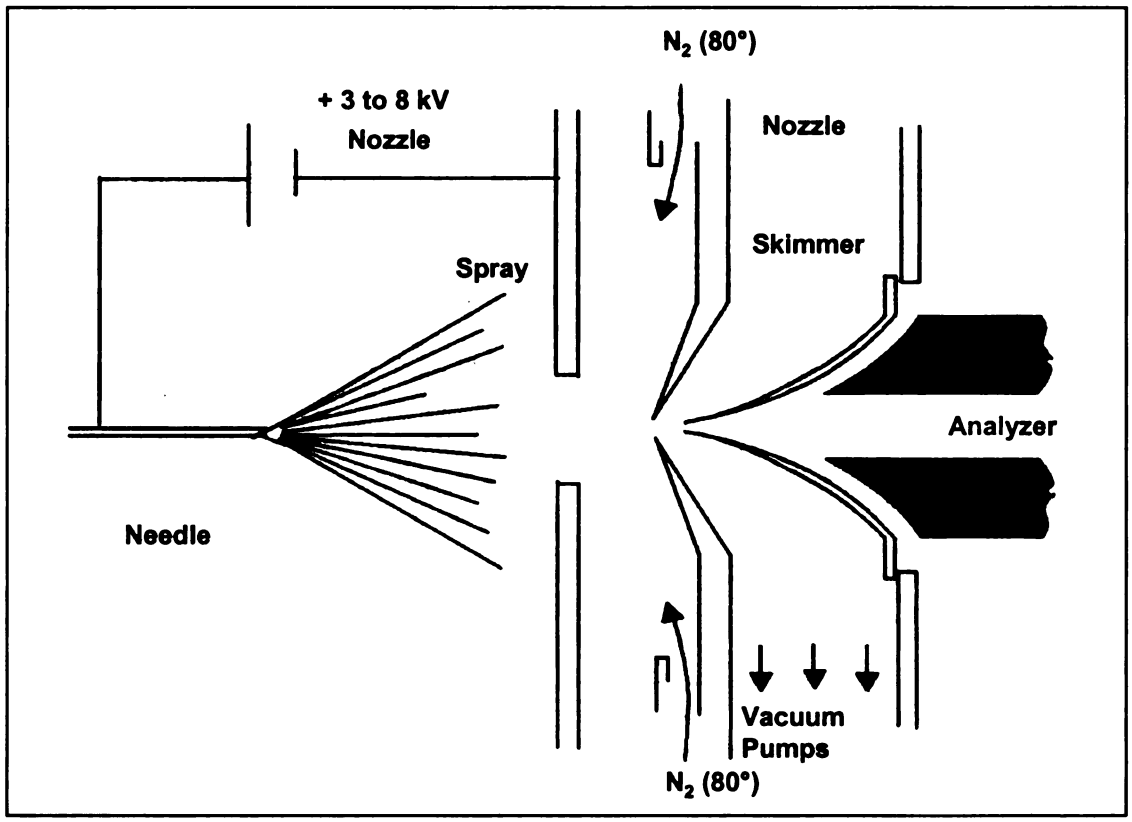


Figure 6. Scheme of an electrospray source (adapted from reference [65])

The emerging liquid at the tube exit starts forming a cone under the influence of increasing field strength. When a certain field strength is reached, the equilibrium between surface tension and electrostatic forces becomes independent of the radius of curvature; theoretically, the radius eventually could become zero. However, in a real system, at the moment the critical electric field

is reached, a conical meniscus that is known as a "Taylor cone" is formed. Immediately, a fine jet of liquid ejects from the cone towards the counter electrode (See Figure 7). The jet carries a large excess of ions because it emerges from the cone's tip that is the point of highest charge density. Such a jet cannot remain stable for an elongated period, and breaks up into small droplets [66].

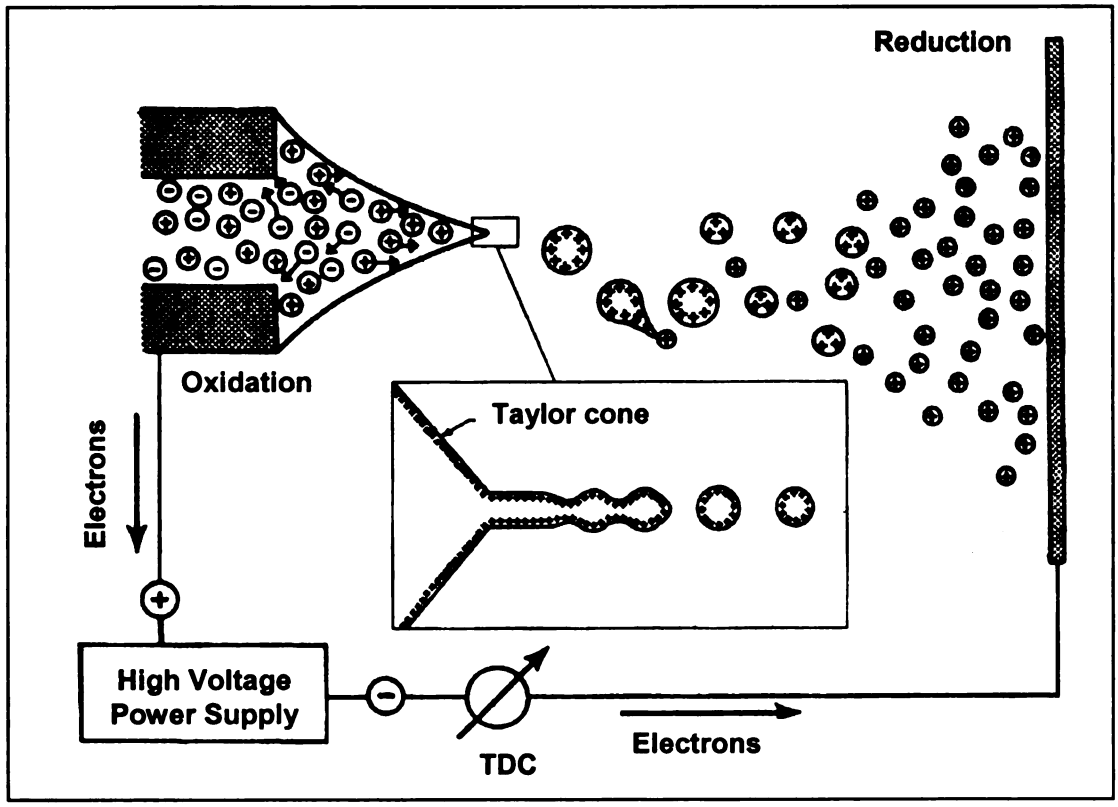


Figure 7. Schematic of Taylor cone formation (from [66]).

The initial fine droplets formed at atmospheric pressure in the spray are enormous at the macromolecular scale; these droplets possess an excess of positive or negative charges depending on the capillary bias polarity. Qualitatively speaking, the charged droplets evaporate to a point where the number of electrostatic charges on the surface becomes so large relative to the

droplet size that an explosion occurs (close to the point that is known as the Rayleigh limit) to produce a number of a smaller droplets. These droplets also contain electrostatic charges [39, 67]. This phenomenon is called Coulomb explosion or Rayleigh instability, and it was captured on film by R. Gomez and Tang in an ESI source (see Figure 8).

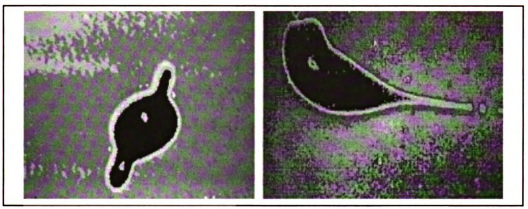


Figure 8. A droplet in the throes of what is known as Rayleigh Instability, and sometimes referred to as Coulomb explosion (from reference [68]).

From a quantitative point of view, for fission to occur, it is necessary that the solvent in the parent droplet evaporate until its radius reaches a value R that leads to the first fission. Droplets close (~80%) and not exactly at the Raleygh limit, with sizes in the 1-μm range, undergo fission because the droplets are deformed mechanically. Typically, the droplets are observed to vibrate alternately from oblate to prolate shapes and these elastic vibrations stimulate disruptions in which the "parent" droplet emits a tail of much smaller offspring droplets (see Figure 8). When relatively volatile solvents such as methanol, acetonitrile, and water are involved and the droplets are a few micrometers in diameter, the evaporation rate follows the surface evaporation limit law, which

leads to a simple dependence of the radius droplet with time *t* (in s), for methanol: R (m) = Ro (m) $- 1.2 \times 10^{-3}$ (m/s) *t* (s), (where Ro is the radio at time zero and R after a time *t*) [69]. Within ~ 100 μs, the radii are sufficiently close to the Raleigh limit where the Coulombic repulsion between the charges overcomes the cohesive force of the surface tension. In this process, a stream of smaller offspring droplets is emitted by the parent. The offspring radius is about one-tenth the parent radius. The offspring carry off 2% of the mass and 15% of the charge of the parent droplet. The decomposition scheme, shown in Figure 9, illustrates the cascade of fissions of a parent droplet. The offspring droplets, on solvent evaporation, undergo Coulomb fissions, which leads to second generation offspring which have only a few elementary charges. However, this decomposition scheme is only partially valid because the data of Gomez and Tang were obtained for larger droplets and different solvents [70, 71].

The next step in the process is the production of gas-phase ions. The mechanism is not completely understood, but there are two main proposed models, both concerning the production of small droplets by means of the Coulomb fissions. However, they have substantial differences. The first model, the Charged Residue Model (CRM) originally proposed by Dole, holds that the sequence of Coulomb fissions produces ultimate droplets, each of which contains only one molecule of solute. That molecule becomes a free gas-phase ion by retaining some of its droplet's charge as the last of its solvent evaporates [72].

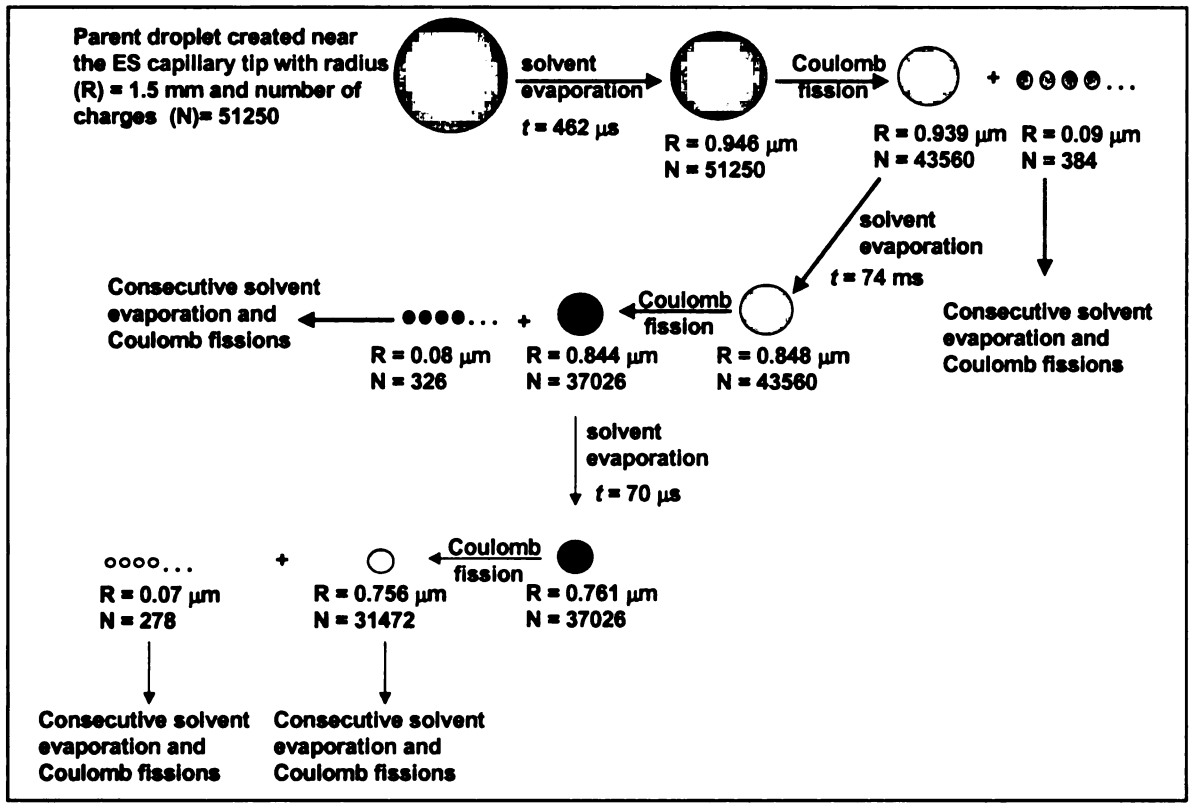


Figure 9. Droplet evolution scheme due to evaporation at constant charge and Coulomb fissions close to the Raleygh limit. Adapted from reference [71].

The second model, the Ion Evaporation Model (IEM) originated by Iribarne and Thomsom [73] assumes the same sequence of evaporation, but argues that before a droplet reaches the ultimate stage contemplated by the CRM ($R \approx 10 - 20$ nm), the field on its surface becomes strong enough to overcome solvation forces and lifts a solute ion from the droplet surface into the ambient gas [72]. Other papers just assert that the gas phase ions produced by ESI are simply the same ions that are in the bulk solution from which the droplet was formed, but that is really a simplistic approach.

Most reported applications of ESI-MS involve molecules in which the charge is most often generated by protonation of basic sites (e. g., amino groups) or deprotonation of acidic sites (e.g., phosphates). Thus, ESI mass spectra in general correspond to a statistical distribution of consecutive peaks characteristic of multiply charged molecules obtained through protonation $(M + zH)^{z+}$ or deprotonation $(M - zH)^{z-}$. Unfolding the proteins, due to reduction of disulfide bonds in a cystinyl protein, could expose additional basic sites to protonation, which results in higher average charge states in the corresponding ESI spectrum. Because these multiply charged species differ only by a consecutively decreasing number of charges as *m/z* increases, it is possible to deduce the mass of the intact molecule. Examples of how to calculate the mass of the intact molecule from ESI-MS spectra are given in detail in textbooks [38, 65-67].

Low-flow electrospray techniques that do not use a make-up solvent have been developed successfully. Microelectrospray, the first low-flow-rate ESI technique, was described using flow rates between 300-800 nL/min [74]. Currently, however, the most common flow-rate technique is nano-ESI; which differs from micro-ESI in the diameter of the needle; flow rates are typically in the range of 20-40 nL/min [64, 75]. The advantages of nano-ESI in comparison to the high-flow rate techniques can be rationalized based on the fundamentals of the ESI process. A lower solution flow rate is known to reduce the size of the charged droplets initially produced in the spraying process. The smaller initial droplets require fewer droplet fission events and less solvent evaporation prior to

ion release into the gas phase. As a consequence, a larger portion of the analyte molecules present in the primary droplets is made available for MS analysis whereas the increase in concentration of contaminants is significantly reduced. Thus, nano-ESI may be considered to be more than just a miniaturized technique [46].

3.2d Atmospheric Pressure Chemical Ionization (APCI)

Like electrospray, APCI creates gas-phase ions at atmospheric pressure. As a consequence, the vacuum and gas dynamics of the interface for APCI are identical to those for electrospray. The sample is sprayed into a heated chamber (~400°C), the ion source. The analyte molecules are volatilized. Nitrogen and oxygen molecules in the ion source are ionized with a corona discharge and these ions react with solvent molecules, which are in the gas phase, to form reagent ions. These reagent ions react with the analyte molecule to produce the analyte ions [39, 76].

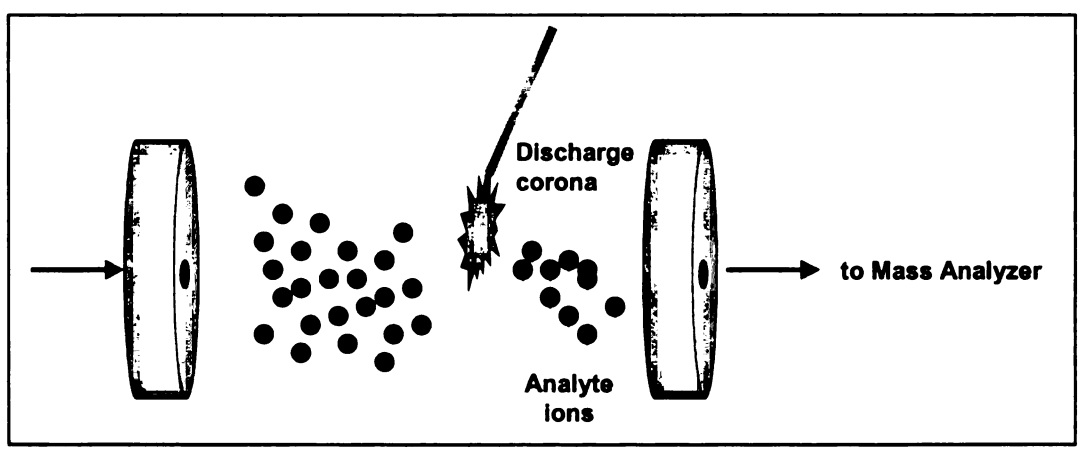


Figure 10. Conceptual representation of Atmospheric Pressure Chemical Ionization.

3.2e Electron Capture Dissociation (ECD)

Electron capture dissociation is a new fragmentation technique used in Fourier transform ion cyclotron resonance mass spectrometry, and it is complementary to traditional tandem mass spectrometry techniques.

This technique was introduced by Zubarev, Kelleher and McLafferty in 1998 [77]. A multiply charged cation captures an electron to produce cationic dissociation products. The cations are generated by electrospray ionization, transferred to the vacuum system of a mass spectrometer, and trapped by electrostatic and magnetic fields in an ion-cyclotron resonance cell. Polypeptide polycations, upon exposure to low-energy thermal electrons (< 0.2 eV), initially capture an electron in a high orbit, which is followed by charge neutralization, leading to an excited radical species that rapidly (10^{-11} s) undergoes bond cleavage. A fraction of ions are reduced to cation-radicals that in part dissociate by cleavage of the N-Cα bonds of the peptide backbone, C_α-alkyl side chain bonds, or cystine S-S bonds [78, 79]. In general terms, the presence of a radical makes this fragmentation very specific, giving cleavages between most of a protein's amino acids.

These days, the mechanism of ECD is poorly understood. Zubarev et al. postulated that this specific and unusual reactivity is due to the fact that ECD is a nonergodic process where the aminoketyl radical dissociation occurs before the

excitation energy (relatively large ~ 6 eV) is randomized, in other words dissociation happens before the reactants undergo internal redistribution of vibrational energy, minimizing the effect of differences in the backbone dissociation energies [77]. This hypothesis for peptide-radical dissociations was based in part on density functional theory calculations in an aminoketyl radical model corresponding to a hydrogen atom adduct to N-methylacetamide, where N-CH_3 bond dissociation was predicted to dominate the spectrum. However, concerns exist about this postulate because non-ergodic dissociation has been observed only for small molecules; thus, other research groups claim that electron capture produces a labile free radical species for which backbone cleavage is the lowest energy reaction [79-81].

From ECD of peptides and proteins it is possible to distinguish different cleavages:

1) Backbone Cleavage: This process produces c (peptide fragment with an enolimine group) and $z^{\bullet}$ fragments (containing an α-radical site) and ~ 10% of $a^{\bullet}$ and y ions [79, 82] (see Figure 11).

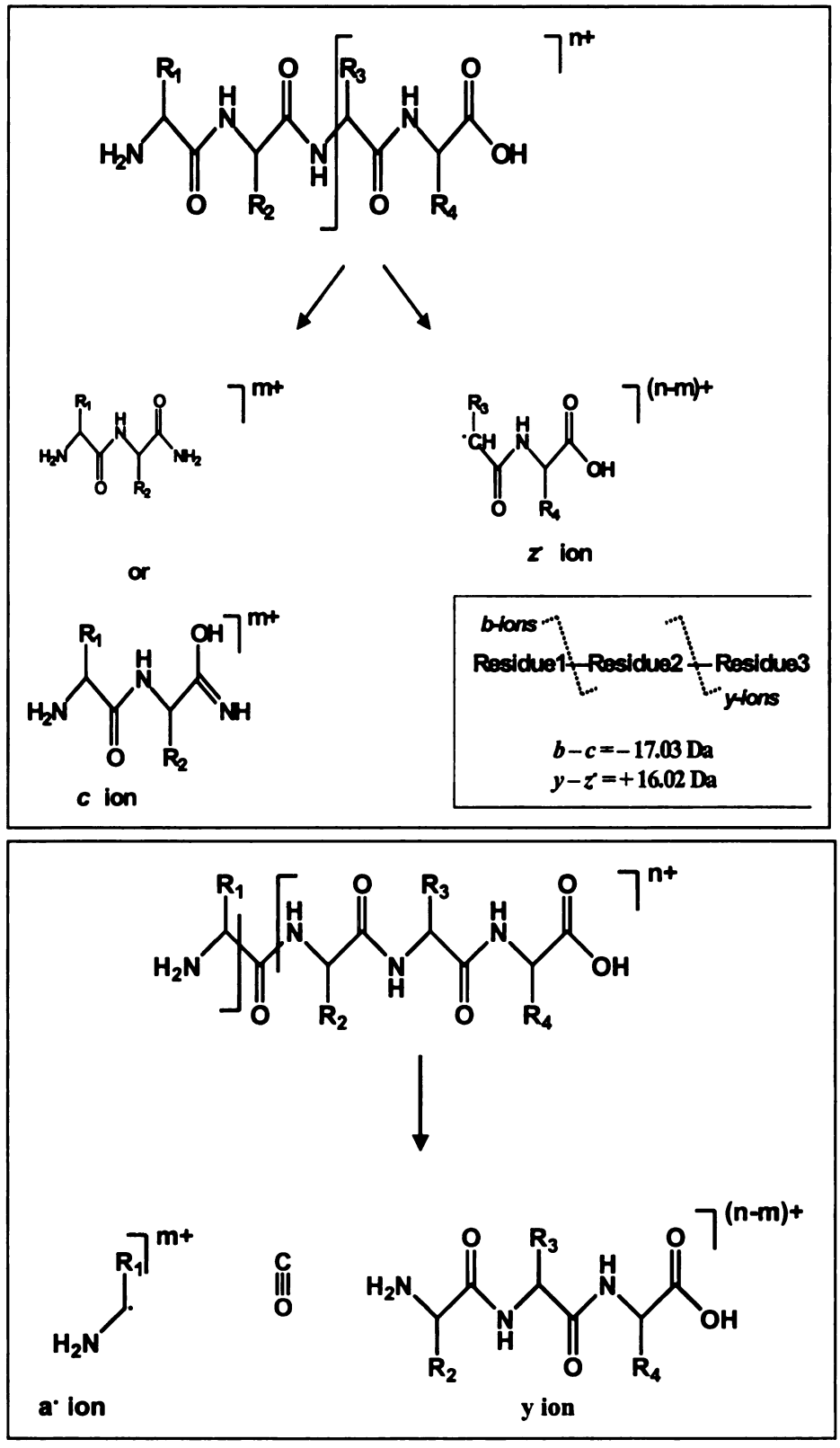


Figure 11. Ions produced by cleavage of the backbone chain in a peptide or protein. Adapted from [82] and [83].

There are no favored sites of cleavage within peptides, except that amine cleavage within proline is rarely observed. The extent of the cleavage in the original ECD seemed to be dependent on the size and conformation of the protein, for example, the examination of the ECD spectra of five 12- to 17-mer peptides showed complete sequence coverage. Backbone cleavage at 16/20 sites was observed for a 21-mer peptide, but only 33 of 152 residue pairs were cleaved in a 17 kDa myoglobin, and none in several larger proteins examined. It has been proposed that the lack of backbone fragmentation in large proteins (> 20 kDa) is due to the presence of non-covalent interactions within the ion, i.e., electron capture results in cleavage of backbone bonds, but does not disrupt the non-covalent interactions. In order to overcome this limitation, "activated ion" (AI) ECD was developed. In this modification, the ions are heated while they undergo electron capture. Activation methods include collisional activation with a background gas, infrared irradiation and blackbody irradiation [83]. AI conditions were sought to denature the center portions of the large ions and help the electron capture dissociation process. Proteins as large as 45 kDa have been analyzed with AI-ECD obtaining acceptable results showing some of the expected ions due to the polypeptide cleavage, and in some cases after optimization, noticeable improvement in the backbone cleavage was found [84-86].

2) Disulfide Bond cleavage: ECD seems to be a good alternative for the analysis of disulfide linkages. The capture of an electron preferentially cleaves

one disulfide bond in proteins. This approach is interesting because intact disulfide bonds can be analyzed in peptides or proteins. [78, 83, 87]

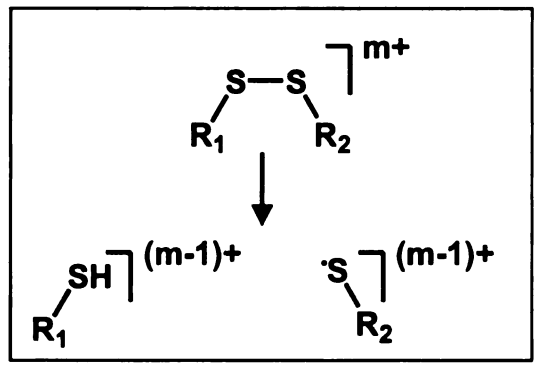


Figure 12. Disulfide bond cleavage using ECD-MS.

3) Side-chain cleavages: These kinds of cleavage are not minor, and in some cases, they show abundances comparable to those coming from backbone cleavages. All peptides containing arginine, histidine, asparagine or glutamine show side-chain losses associated with those residues, regardless of the residue position within the peptide. For example, arginine side-chain cleavages result in losses of CH_4N_2 (44.037 Da), CH_5N_3 (59.048 Da), and $C_4H_{11}N_3$ (101.095 Da). Losses associated with lysine and methionine were observed, but because these losses have not been found in all peptides containing those residues, so the losses cannot be considered of absolute diagnostic value. Other side-chain cleavages with possible diagnostic value have been reported, but currently it is quite early to know if those losses are always present in the ECD process.

3.3 Mass analyzers

Ionized matter can be affected by electric and magnetic fields. Thus, after matter is ionized in an ionization source, ions are subjected to electric and magnetic fields in order to separate them. Acceleration, energy and momentum induced by magnetic and electric fields are dependent on the charge on the ion, thus mass analyzers separate ions according to mass-charge (*m/z*) ratio, not according to their mass only.

In principle, all ionization techniques could be coupled to any mass analyzer, but some couplings work better than others because of the particular characteristics of the mass analyzers, the ionization techniques, and the type of information under investigation. Mass analyzers have some common characteristics: 1) mass range (maximum *m/z* detectable), 2) resolving power (ability to separate ions of adjacent *m/z)*, 3) accuracy of mass measurement, and 4) ion transmission, sensitivity [38, 88].

Next, there are brief descriptions about mass analyzers commonly found in the analysis of biomolecules.

3.3a Quadrupole Mass Analyzer (Q)

This instrument separates ions based on radiofrequency (rf) oscillations in an electric field (the quadrupole field) superimposed on direct current (DC) potentials. This mass spectrometer employs a combination of DC and rf to filter out all ions except those of a given *m/z* value at a given set of rf and DC amplitudes. A scan of the amplitudes of rf and potentials allow ions of a sequential range of *m/z* values to reach the detector to generate a complete mass spectrum.

Quadrupoles consist of four metal rods held in strict alignment with one another. Opposite rods are connected in pairs to both radio frequency and direct current potentials. According Figure 13, ions coming form an ionization chamber are accelerated at the entrance of the quadrupole by means of an acceleration potential in the Z direction. Within the quadrupole, ions experience no forces in the Z axis, however, ions traveling along the z-axis are subjected to a total transverse electric field resulting from the application of the potentials upon the rods:

$$\phi_o = + (U - V \cos \omega t) \qquad \text{and} \qquad -\phi_o = - (U - V \cos \omega t)$$

ϕ_o represents the potential applied to the rods, ω is the angular frequency (in rad/s) = $2\pi \nu$, where ν is the frequency of the rf, U is direct potential and V is the "zero to peak" amplitude of the rf voltage [38, 67, 89, 90].

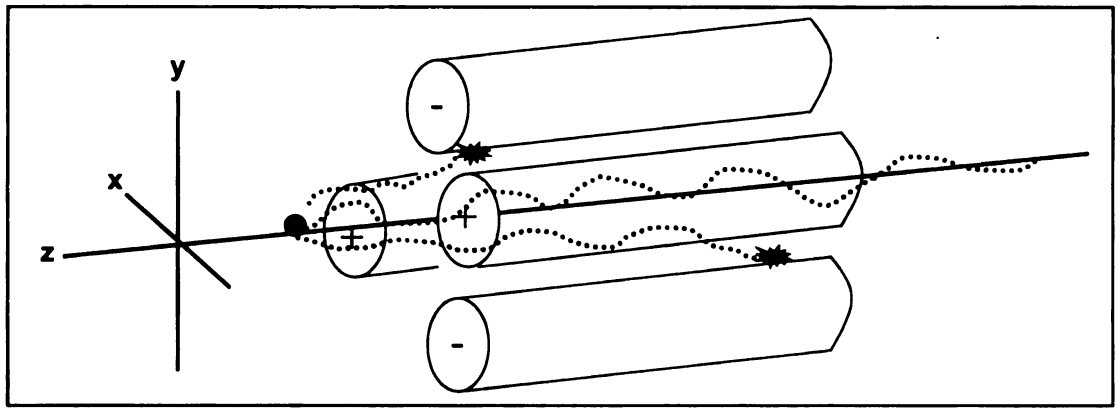


Figure 13. Quadrupole mass filter.

Depending of the potential applied, ions within a small range of *m/z* values have stable paths through the quadrupole and all others ions are not transmitted. Ideally, for specific values for V and U, only ions of one *m/z* value should be able to have stable trajectories along the X-axis and along the Y-axis, thus those ions could travel in the Z-direction and by consequence, be able to reach the detector after exiting of the quadrupole mass filter. On the other hand, ions with unstable trajectories either in the X-axis or Y-axis, will strike the rods in the quadrupole, become discharged and not detected.

Equations of motion for an ion can be depicted in the form of the Mathieu equation (which was established in 1866 by the French physicist E. Mathieu):

$$\frac{d^2u}{d\xi^2} + (a_u - 2q_u \cos 2\xi)u = 0$$

where *u* stands for either x or y. These equations are quite complex and derivation implies solution of differential equations that produce a set of variables a, q, and ξ:

$$a_u = a_x = -a_y \frac{8U}{\omega^2 r_o^2}\left(\frac{z}{m}\right) \quad q_u = q_x = -q_y \frac{4V}{\omega^2 r_o^2}\left(\frac{z}{m}\right) \quad \xi = \frac{\omega t}{2}$$

From these variables, we can see that it is possible to establish a relationship between the coordinates of an ion and time using the Mathieu equation; detailed integration and analysis of the motion equations can be found elsewhere [88, 91, 92]. For ions of a given mass, certain values of *a* and *q* lead to stable oscillations in the x and y directions. For a given quadrupole, r_o is constant, ω is maintained constant and U and V are the variables. Since *a*, *q*, *U* and *V* are interrelated, *m/z*, varies with either *U* or *V*. The relationship between all of these variables and the determination of *m/z* can be understood in terms of a plot of *a* vs *q* (See Figure 14).

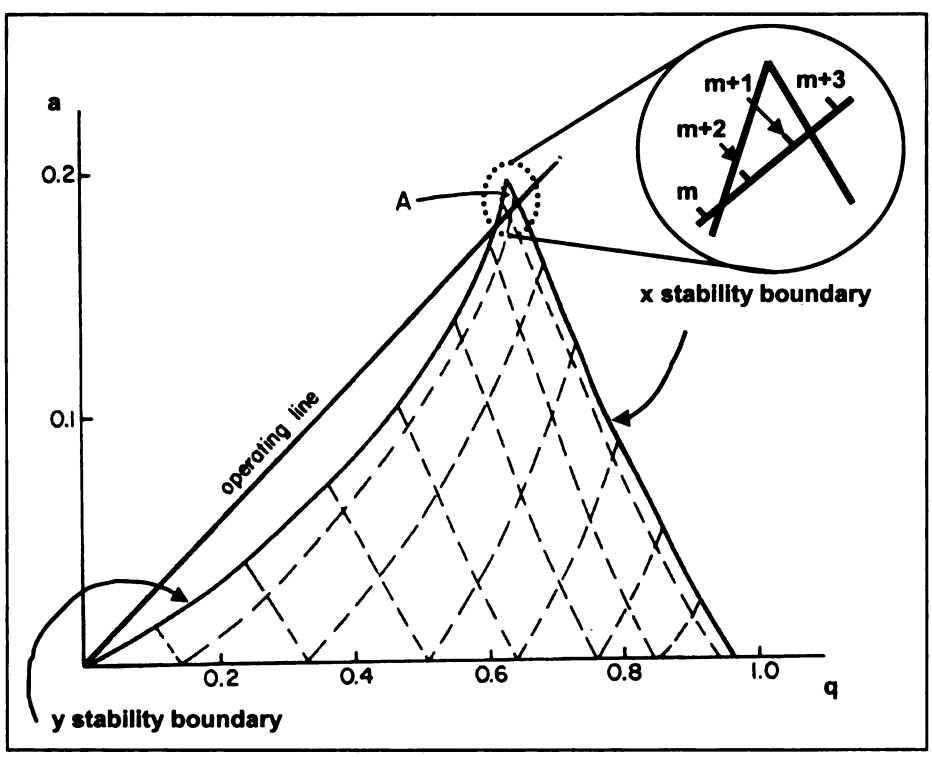


Figure 14. Quadrupole stability diagram. (Adapted from [38, 88, 92])

This graph is the collection of points in *a-q* space that correspond to stable solutions of the equation of motion (shadowed area in Figure 14) in physical terms; this diagram is the reduction of a six-dimensional problem (involving *z*, *m*, U, V, r_o, and ω) to a two-dimensional problem involving the reduced parameters *a* and *q*.

We can describe the operation of a quadrupole as a narrow bandpass mass filter or as rf-mass filter in terms of the a-q stability diagram.

In the first case, let's consider that, in principle, it is possible to operate a quadrupole in a manner in which the parameters *a* and *q* are, at all times, independent from one another. In practice, quadrupoles are operated in a manner that the values of those parameters are always related by a simple ratio regardless of the actual magnitude of either *a* or *q*:

$$\frac{a}{q} = \frac{2U}{V}$$

In terms of the stability diagram, the above equation represents a straight line with a zero intercept that is called operating line (see Figure 14). This line can be visualized from a practical point of view like a collection of different masses that might be stable if they are inside the stabilization diagram, for example, in the diagram, masses m+1 and m+2 are stable (they are inside the region A), on the contrary; masses m and m+3 are not stable. If the slope of the operating line is increased, the U/V ratio is higher, the area A will become smaller

and it would represent the case in which only a few masses (maybe only m+1) will have a stable trajectory, meaning that the resolution of the quadrupole is increased. From equations for *a* and *q*, it is known that U $\propto$ (*m/z*)a and V $\propto$ (*m/z*)q, then if U or V are increased (keeping constant ω and r_o), *m/z* values should be increased also. Thus, the most convenient method of scanning the bandpass region is by sweeping the voltage applied; consider that if potentials U and V are doubled, the masses that appear in the area A must be doubled, therefore, a potential increase is equivalent to sliding the operation line. The complete mass spectrum is obtained by scanning the magnitude of V and U through increasing or decreasing values; by scanning the magnitude of rf and DC potential from a low to high value, ions of increasing mass will sequentially take on stable values of *a* and *q* space and be transmitted to the detector. For example, for ions in Figure 14, the order would be: m $\rightarrow$ m+1 $\rightarrow$ m+2 $\rightarrow$ m+3.

For the case of using the quadruople in the rf-only mode, the stabilization diagram is very useful, because this operation mode is represented when U = 0, it means a = 0 (see Figure 15).

We can observe in this way that a large portion of the operating line falls within the stability region; it predicts that a large number of ions with different *m/z* values have stable trajectories within the quadrupole. Ions of all *m/z* ratios higher than a particular lower limit will have a stable trajectory (in this case, higher than m-1). In this mode, the quadrupole causes all ions to be brought

back systematically to the center of the rods, even if they were deflected by a collision. This focusing effect is important to increase the transmission of ions after collisions [65, 67, 91, 92].

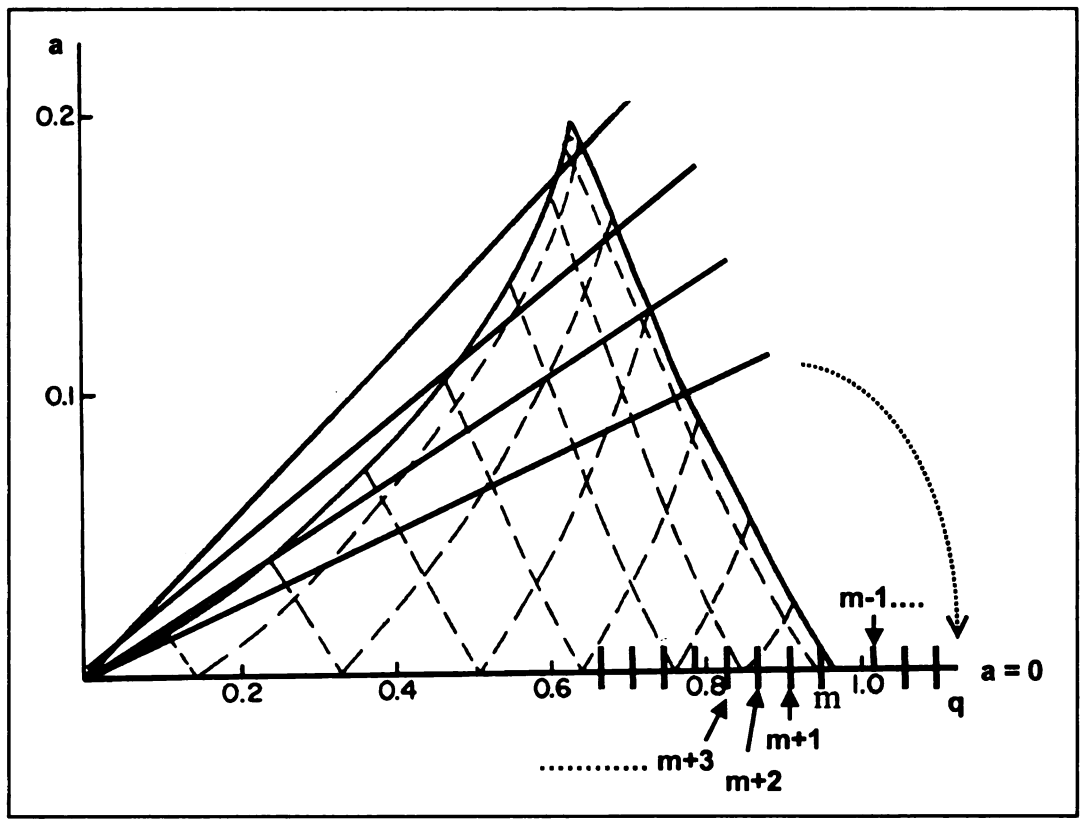


Figure 15. Quadrupole stability diagram when working in rf-only mode (a = 0). (Adapted from [38, 88, 92])

Quadrupole mass spectrometers are symbolized by "Q" and rf-only quadrupoles by "q"; these letters are commonly used in hybrid mass spectrometers where quadrupoles are an essential part of the instrumentation setup. Quadrupole-based instruments are among the most sophisticated instruments for structure elucidation, and for this reason, it is necessary to have a good understanding of the theory associated with the design and operation of quadrupoles [91].

3.3b Quadrupole Ion-Trap (QIT) or Quistor (Quadrupole Ion Storage)

The invention of the ion trap (in 1953) by Wolfgang Paul and Hans Steinwedel, was recognized by the award of the 1989 Nobel Prize in Physics. In 1983, the mass-selective instability scan by George C. Stafford, Jr. was announced. On these two landmarks rests the entire field of ion trap mass spectrometry [93].

In the quadrupole ion trap, the ions are trapped in a small volume between a ring electrode and two end-cap electrodes in a three-dimensional space. Like the quadrupole mass filter, it uses oscillating electric fields (rf) to trap ions in a controlled manner. The common operation mode of an QIT is with the potential applied to the ring, while the end-cap electrodes are grounded except when auxiliary potentials are applied to either or both end-cap electrodes [94].

In quadrupole instruments, the potentials are adjusted so that only ions with a selected mass go through the rods, in the case of QIT, ions within a range of *m/z* values are stored in a field that is created with a fixed-frequency rf applied to the cylindrical ring electrode (see Figure 16). As the amplitude of this fixed-frequency rf is increased, the trajectories of ions on increasing *m/z* values will become sequentially unstable and move toward the exit end cap electrode, which has holes to allow ions to leave the QIT, and go to the detector to be recorded [39, 94].

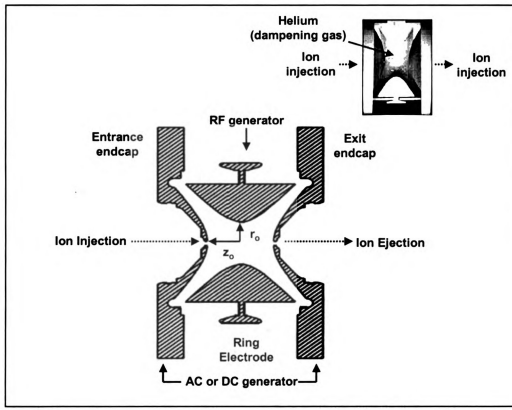


Figure 16. Schematic diagrams of an ion-trap; z_o is distance from the center of trap to apex of the end-cap, r_o is the internal radius of the ring electrode. (Adapted from [93]).

There are two types of QITs, 1) instruments where the primary ionization takes place in the space where ions are stored, 2) instruments where the primary ionization takes place outside the QIT; these last types of QITs are currently the most common in mass spectrometry instrumentation applied to biomolecules.

The mathematical analyses using the Mathieu equations allow us to locate areas wherein ions of given masses have a stable trajectory. The equations are very similar to those used for quadrupoles, however, in the ion trap, the motion of the ions occurs in three dimensions (see A in Figure 17).

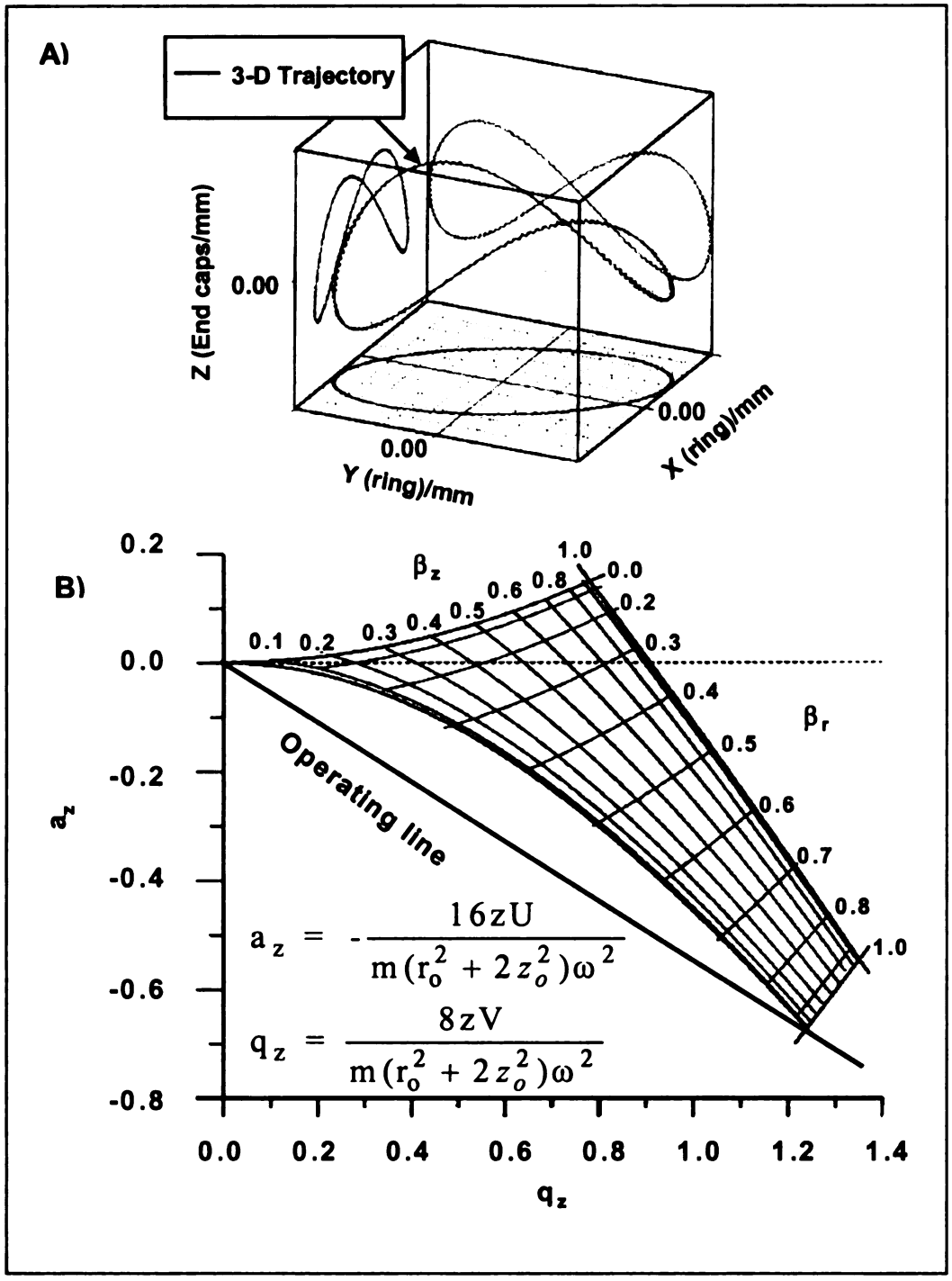


Figure 17. A) Simulation in two- and three- dimensions of the trajectory of an ion with a *m/z* value equal to 500 (Diagram from reference [94]). B) Stability diagram for the ion-trap, subscript z represent axial motion between the endcaps. U is the DC potential on the endcap electrodes, V is the rf potential applied to the ring electrode, r_o is the radius of the ring electrode, ω is the rf angular frequency, z is the charge and m is the mass. Parameters β are additional trapping parameter to q and a, to describe the ion stability in the diagram (From reference [95]).

In the stability diagram presented in Figure 17, it is interesting to note that the stability region is only a small area in the *a-q* field. The parameter β is an additional trapping parameter to *q* and *a* to describe the ion stability in the diagram. It comes from the complete integration of the Mathieu equation which requires the use of a function $e^{(\alpha+i\beta)}$. The behavior of an ion can be depicted by means of the parameter β value. As the value of β approaches to zero, the ion is not affected by the oscillating fields; when β is equal to 1, the ion oscillates in resonance with the rf field and absorbs kinetic energy to the point that its magnitude of oscillation increases so that the ion escapes the trap or is discharged by collision with the end cap surfaces. For β values between 0 and 1, the ion oscillates in a periodic mode caused by the oscillating fields with an average kinetic energy related to the value of β [38].

A relatively high pressure (10^{-2}-10^{-3} Torr) of inert gas, usually He, is used to dampen the motions of the ions, confining their coordinates to the center of the trap. Under normal operation, ions will have a sufficient β value so as not to have enough kinetic energy to fragment upon collision with the inert gas. Increasing the value of β for specific ions, it is possible to establish appropriate conditions for ion/molecule experiments or collisionally activated dissociation (CAD or sometimes called CID, from collision-induced dissociation).

It is possible to distinguish four different ion traps functions:

Collection: Ions are injected and cooled by means of the inert gas in the center of the trap.

Isolation: For tandem experiments (MS^n) or for Selected Ion Monitoring (SIM), a selected ion of a *m/z* value is stored in the trap, all the others ions are ejected from the trap.

Excitation: For tandem experiments, after ions of one specific *m/z* value are isolated, they can be excited and collided with the inert gas to obtain product ions.

Ejection: Ions in the trap are ejected and detected to get the mass spectrum.

These functions are programmed depending on the desired experiment.

There are two main drawbacks for the QIT, 1) Limited storage capacity and 2) Inefficient trapping. In the first case, there is a maximum number of ions that can be stored within a trap at any one time. When the QIT is filled beyond this maximum, "space charge" occurs. This phenomenon is observed when repulsive forces cause leakage of the excess ions in a trap. Concentration of ions above the space-charge limit leads to poor performance of the instrument, lowering resolution and sensitivity, and sometimes affecting calibration (mass accuracy). In a 3D trap, the presence of a 3D rf field is unfavorable for ion

trapping, for example, during the positive cycle of the field, positively charged ions are repelled from the trap by the positive electrical field, and only when this field is negative, positively charged ions are pulled through the trap. However, the possibility of isolation and fragmentation of specific ions (tandem in time mass spectrometry) make the QIT a versatile tool of ample application in mass spectrometry [38, 39, 93].

3.3c Linear (Quadrupole) Ion Trap (LIT)

Linear ion traps have found new applications in mass spectrometry recently, either combined with other mass analyzers such as time-of-flight (TOF), quadrupole ion traps (QITs), or ion cyclotron resonance (ICR), or as stand-alone mass spectrometers with radial or axial mass-selective ejection. In a LIT, ions are confined radially by a two-dimensional radiofrequency field and axially by stopping potentials applied to end electrodes. In comparison to QIT, LIT have higher injection efficiencies and higher ion storage capacities. Their use is not limited to simply storing ions, they can be operated as stand-alone mass spectrometers [96].

Many of the techniques used to manipulate QITs can be applied to linear quadrupole traps. These include resonant ion excitation for ion activation, isolation, or ejection, and mass-selective instability scans with ejection radially through a rod (or rods, see slot in Figure 18) or axially.

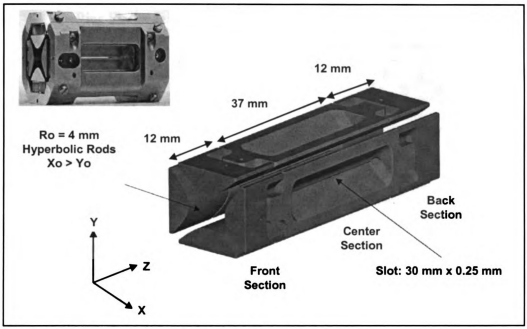


Figure 18. Linear Ion Trap from Thermo-Finnigan [93].

The advantages associated with trapping ions in a LIT are several-fold, because they have no quadrupole field along the z-axis. Ions admitted into a pressurized LIT undergo a series of momentum-dissipating collisions effectively reducing axial energy prior encountering the end electrodes thereby enhancing trapping efficiency. The larger volume of the LIT relative to a QIT means that more ions can be stored prior to undergoing charge space effects. Radial containment of ions within a LIT results in strong focusing along the centerline of the trap in contrast to the QIT in which the fields tend to focus the trapped ions to a point (see Figure 19). Line rather than point-focusing properties may have some influence on the relative susceptibilities to space charge phenomena [97], in addition, conventional 3-D ion traps have a slower scan rate.

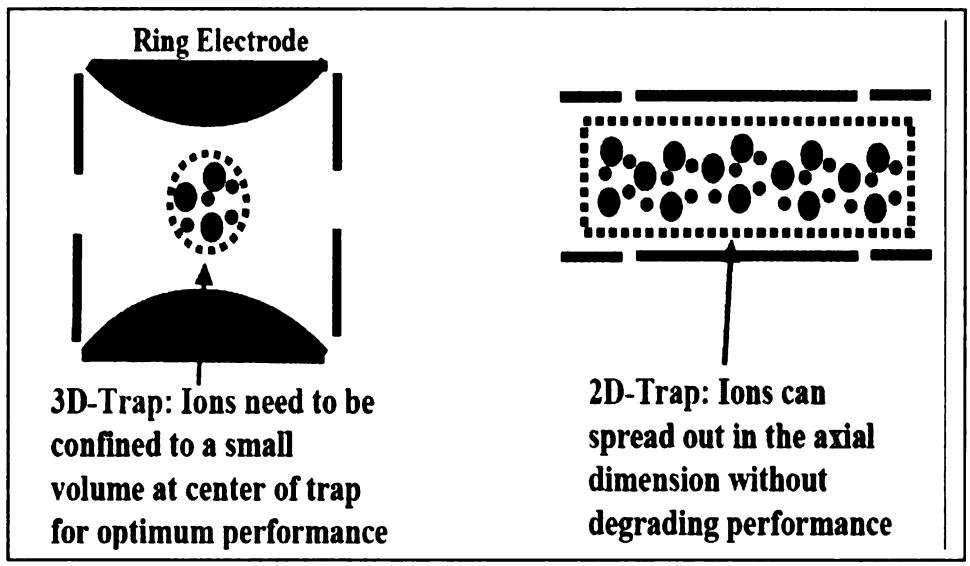


Figure 19. Comparison of focusing in a QIT and a LIT, from reference [93].

These problems with 3D ion-traps may contribute toward a reduced number of peptides identified as well as limited quality fragmentation spectra during proteomics experiments. Functional enhancements of the LIT include 15x higher ion capacity, 3x faster scan rate, up to 100% detection efficiency, and up to 70% trapping efficiency.

In a LIT, mass-selective radial ion ejection involves trapping the ions within the four-rod quadrupole structure and resonantly ejecting the ions radially through a slot cut in one of the quadrupole rods. In the commercial LTQ (from Finnigan), the ions are moved from within the stability diagram to a position where they become unstable in the x-direction and leave the trapping field for detection through double slot cuts; from there, ions are accelerated into two high voltage dynodes where ions produce secondary electrons. This signal is subsequently amplified by two electron multipliers; the analog signals are then

integrated and digitized. This dual detection helps to increase sensitivity in the process.

Mass-selective ejection from a linear trap is a less obvious technique for ion extraction, but it is analogous to the rf-only transmission in a quadrupole mass spectrometer. Because axial ejection linear ion trap operates in the low 10^{-5} torr regime and ions emerge from the end of the device; it has proven to be a good match with the ion path of a QqQ instrument [98].

Most other features and operations of LITs are similar to those of the QITs, and have been described in detail in several articles [96-101]. A systematic comparison of LITs and QITs in proteomics experiments found a 4-6 fold increase in the number of peptides and proteins identified on the LIT mass spectrometer and, interestingly, more than 70% of the double and triply charged peptides, but not singly charged peptides, showed better quality of fragmentation spectra on the two-dimensional ion trap [98].

3.3d Time of Flight (TOF)

This instrument uses no external force to separate ions of different *m/z* values. Ions are accelerated into a flight tube at values from hundreds to thousands of volts (typically 20 kV). The operating principle of a TOF mass spectrometer involves measuring the time required for an ion to travel from an ion source to a detector located 1-2 m from the source. All ions receive the same kinetic energy during instantaneous acceleration (e. g., 3000 eV), but because they have different *m/z* values, they separate into groups according to velocity (and hence *m/z*). The ions sequentially strike the detector in order of increasing *m/z* value.

The simplicity of the physical principles of TOFMS is part of its current success. The mass-velocity relationship for constant energy ions and the electrostatic force on a charge when combined with the Newtonian mechanics provide the foundations of all TOF mass spectrometers:

$$qV = \frac{1}{2}mu^2$$

$$F = Eq$$

$$F = ma$$

$$a = \frac{Eq}{m}$$

where:

q = charge (Coulombs)

V = electrical potential

m = ion mass

u = ion velocity

F = force experienced by charge

E = electric field strength

a = acceleration

An ion with mass m and total charge q = ze has a kinetic energy:

$$E = \frac{mu^2}{2} = qV = zeV$$

e = charge of an electron

z = charge state of the ion

The time (t) needed to fly the distance (d) is giving by:

$$t = \frac{d}{u}$$

Thus,

$$\frac{m(d/t)^2}{2} = qV = zeV$$

$$\frac{m}{2}(d/t)^2 = zeV$$

$$\frac{m\,d^2}{2} = zeVt^2$$

$$t^2 = \left(\frac{m}{z}\right)\frac{d^2}{2Ve}$$

This equation shows that *m/z* can be calculated from a measurement of *t*, because *d*, and *V* are kept constant. As example, the transit time at 3000 eV of an ion of *m/z* 800 is approximately 70 μsec through a 1-m flight tube [38, 67, 102].

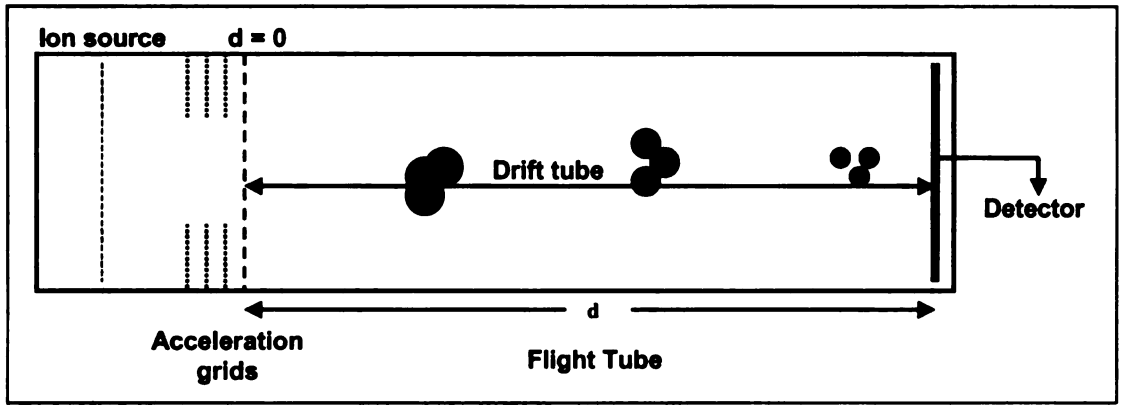


Figure 20. Scheme of a Time Of Flight (TOF) mass spectrometer.

In principle, the upper mass range for a linear TOF has no limit. The most important drawback of a linear TOF analyzer is its low resolving power; four basic phenomena are responsible for this (see Figure 21) [76]:

1) Temporal distribution: Two ions of the same mass formed at different times in the source, thus, they arrive at the detector at different times.

2) Spatial distribution: Two ions of the same mass formed at different locations in the source arrive at the detector at different times.

3) Kinetic energy distribution: Two ions of the same mass formed with different initial kinetic energy arrive at the detector at different times.

4) Turn-around time effect: Two ions of the same mass with the same kinetic energy, but initial velocities in opposite directions arrive at the detector at different times

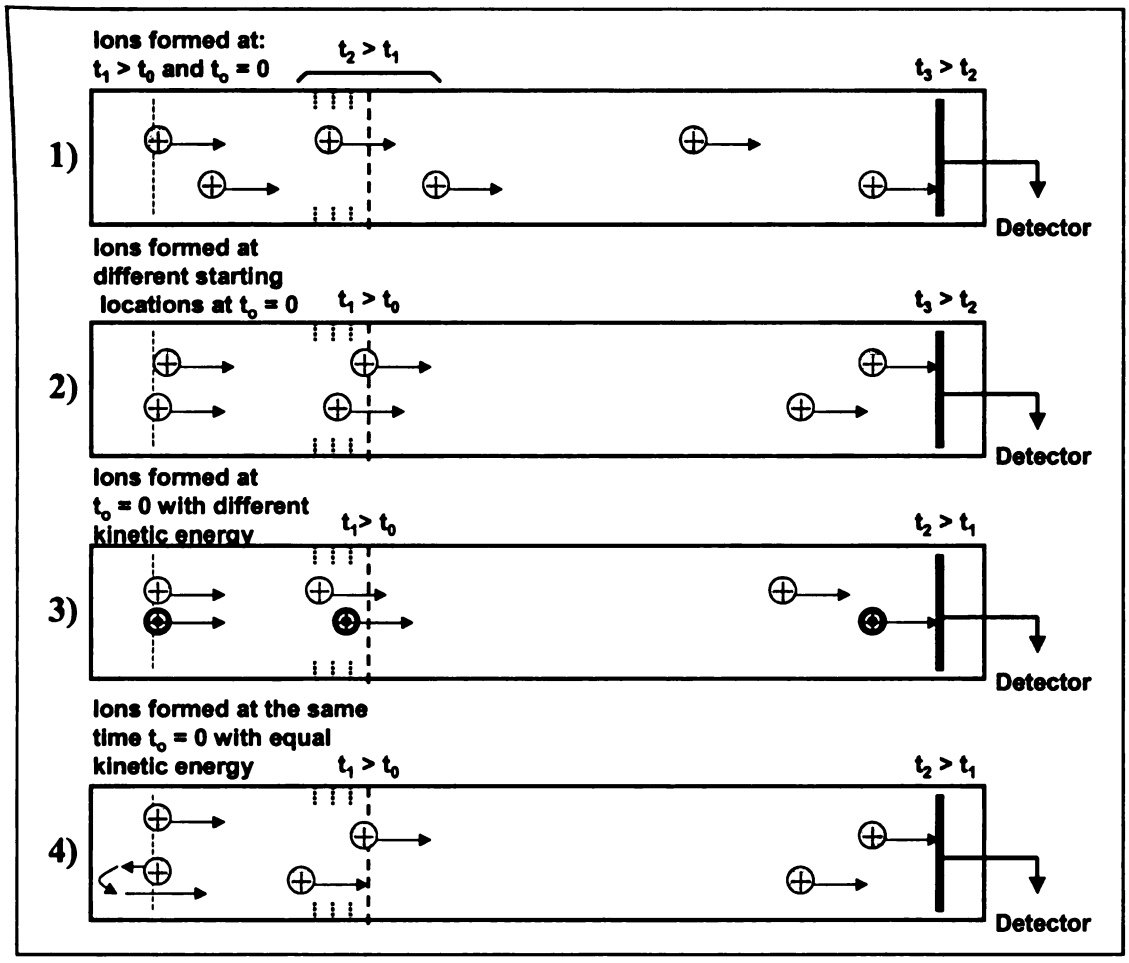


Figure 21. Phenomena that affect resolving power in a TOF mass spectrometer. 1) Temporal distribution 2) Spatial Distribution 3) Kinetic energy distribution 4) Turn around time effect.

Several techniques have been developed to correct these problems, spatial variations were improved by a unique two-field extraction and acceleration source, energy variations and turn-around time were corrected by a unique technique called time-lag focusing. These techniques can be found nicely described in some textbooks [38, 66, 67] and articles [102-106]. Another major improvement in resolving power involves the use of an electrostatic reflector, also called a reflectron that was proposed by Mamyrim and Shmikk [107]. Let's considerer that isomass ions, with a distribution of energies, travel at different

velocities. After traversing the field-free region, they reach a series of grids and ring electrodes of increasing potential (the reflectron) that create a retarding field that acts as a mirror by deflecting the ions and sending them back through the field tube; indeed, ions with more kinetic energy penetrate more deeply into the reflector, spending more time, therefore they will reach the detector at the same time as slower ions (see Figure 22). However, sensitivity of this mode of operation is low in comparison to the usual linear mode [38, 66, 67].

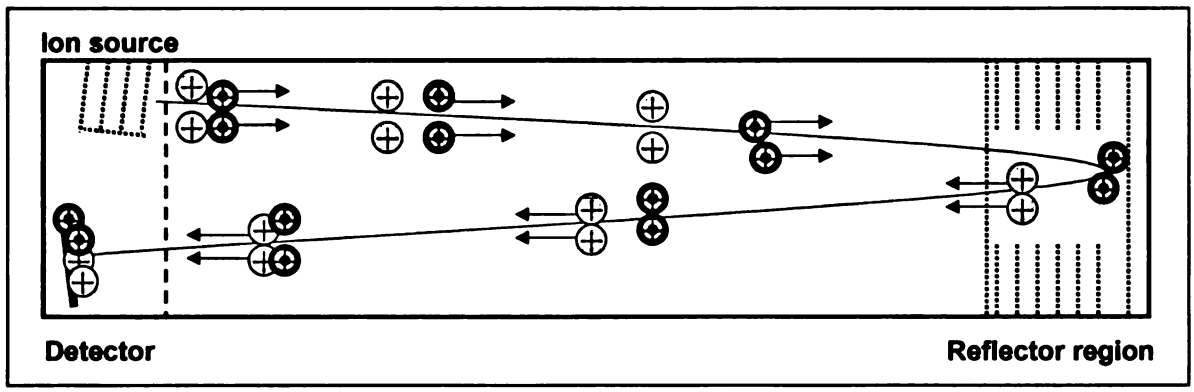


Figure 22. Schematic diagram of ion trajectories in a reflectron-Time-of- Flight mass analyzer.

A major breakthrough has been the development of an orthogonal acceleration TOF (oa-TOF) analyzer, where "turn-around time" is not a problem. In this mass analyzer, pulsed ions are extracted orthogonally (perpendicular) from a continuous ion beam [66, 104, 106, 108]. Ions are sampled from a nearly parallel ion beam (see Figure 23), the low-energy ion beam is allowed to fill an ion accelerator region while it is held in a field-free state (in x-axis). Then, a package of ions is pushed out orthogonally (in y-direction) by means of a sharp pulse and it is accelerated into the TOF analyzer (y-direction). As the beam is nearly parallel, the ions have nearly zero average velocity and minimal velocity

spread in the direction of the pulse applied. After the heaviest ions have reached the detector, the next package is pulsed into the TOF-analyzer (Figure 23).

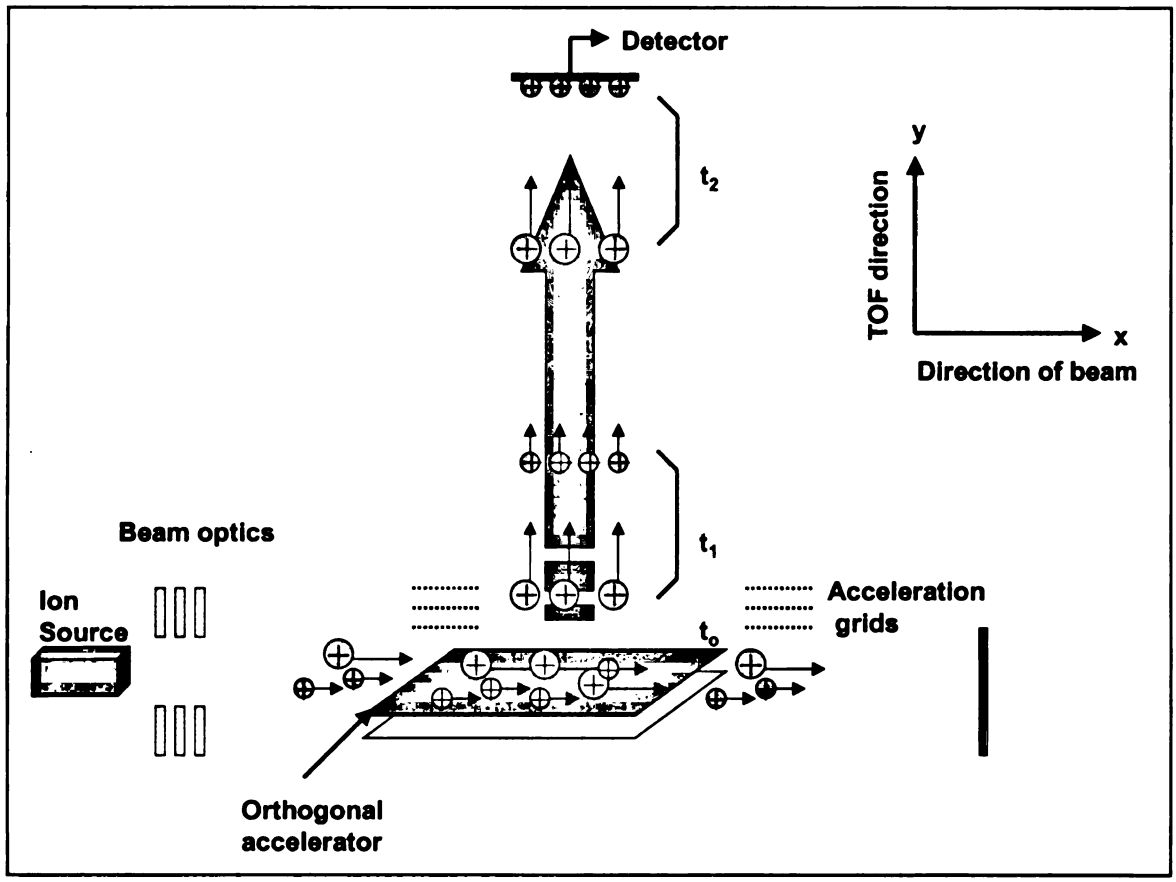


Figure 23. Scheme of a linear oa-TOF. t_o, t_1 and t_2 represent the flight time of ions after the orthogonal pulse.

Advantages of oa-TOF are: 1) high sensitivity, 2) high rate of spectra per second, 3) high mass-resolving power, 4) mass accuracies up to 5 ppm, and 5) compact design [66, 104, 106].

3.3e Fourier Transform Ion Cyclotron Resonance (FT-ICR)

The Fourier transform ion cyclotron resonance (ICR) mass spectrometer is a magnetic ion trap. The analyzer employs a powerful magnetic field (3-7 Tesla). Recall from physics that a charged particle precesses in a magnetic field at a frequency inversely proportional to its mass, this motion is known here as the "cyclotron motion". Formerly, an ion of velocity v and charge q (where q = ze, e = 1.6021 x 10^{-19} C), entering into a uniform magnetic field B perpendicular to its direction will move on a circular path of radius r (see Figure 24), determined by the equation:

$$r = \frac{mv}{qB}$$

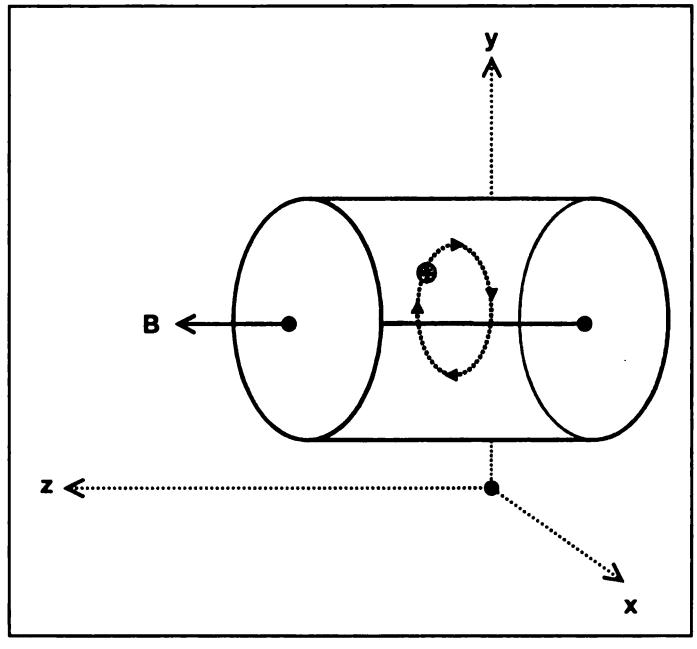


Figure 24. Precessing of an ion in a magnetic field.

the angular frequency ω (in rad/s) about the z-axis is defined by:

$$\omega = \frac{v}{r} = 2\pi f$$

then: $$\omega = \frac{qB}{m} = \frac{zeB}{m} = \frac{zeB}{m/z} = 2\pi f$$

$$f = \frac{eB}{2\pi m/z}$$

where: f = frequency (Hertz, Hz)

B = magnetic field (Tesla, T)

z = multiples of elementary charge

m = mass in Da, ($m = m_{(\text{in kg})} \times \frac{1\ \text{Da}}{1.66054 \times 10^{-27}\ \text{kg}}$)

From the equations above, we can notice that: 1) the cyclotron angular frequency (ω) is independent of the ion's initial velocity, but a function of its mass, charge, and the magnetic field, 2) the frequency is also independent of velocity, but depends on the ratio (q/m)B, 3) the radius of the trajectory increases, for a given ion, proportionally to the velocity, thus if the radius becomes larger than that of the cell where the ion is confined, the ion is expelled [38, 66, 67, 109, 110].

A Fourier Transform Ion Cyclotron Resonance Mass Spectometer is also an "ion trap". It consists of a cubic cell (the trap) inside a strong magnetic field.

The cell has three distinct sets of plates: trapping, transmitter, and receiver plates. For mass analysis, a continuous beam of ions is formed outside the cell and pulsed into it. When ions enter the FT-ICR analyzer cell, they are constrained by the strong magnetic field such that they have a cyclotron motion; ions are also constrained by electric potentials applied to a set of trapping plates perpendicular to the magnetic field (see Figure 25) [76].

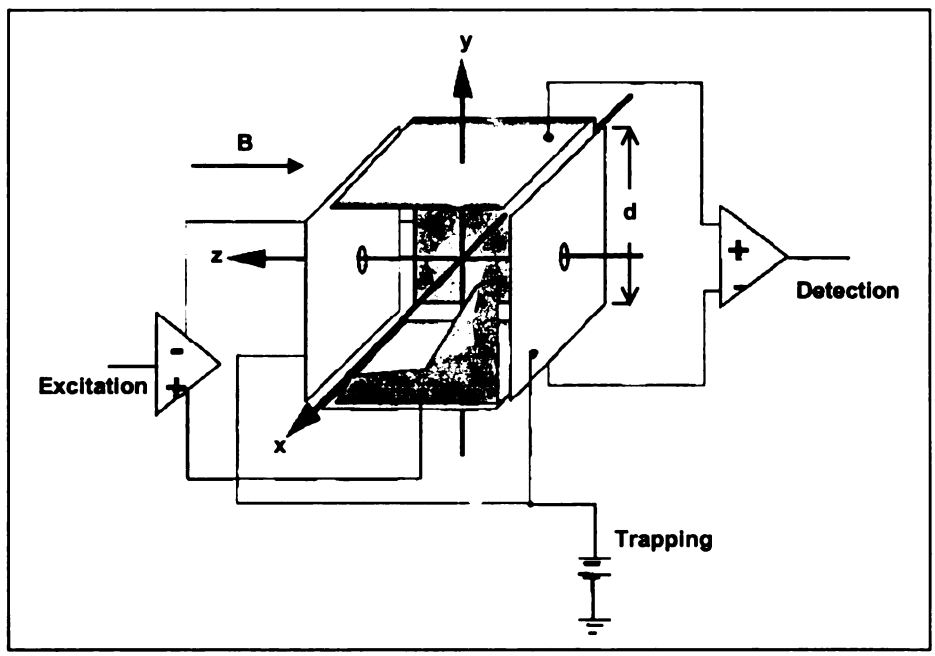


Figure 25. Schematic diagram of a FT-ICR mass analyzer (adapted from [110]).

Thermal energy produces a circular micromotion of the ions in the xy-plane as they enter to the magnetic field; however, the radii of these orbits are too small to be detectable. Ions of each mass have their characteristic cyclotronic frequency. It can be demonstrated that ions excited by AC irradiation at their own frequency and with the same energy [the same potential (V_{p-p}/d),

applied over the same time, T_{exc}] will have an orbit with the same radius, in other words, the orbital radius is independent of *m/z*: $$r = \frac{V_{p\text{-}p} T_{exc}}{2dB}$$

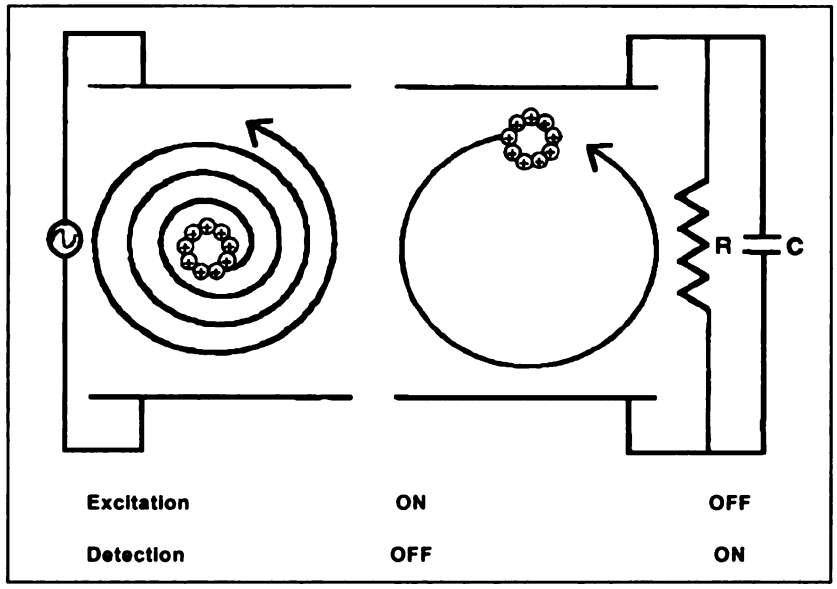


Figure 26. Sequence of excitation (left) and detection in FT-ICR-MS. The ionic micromotion is indicated by the small circles of ions. (Adapted from [66, 110])

In summary, all ions present in the cell are accelerated with an rf sweep (commonly called a chirp) that covers a range of frequencies corresponding to the *m/z* values we wish to observe; after excitation, ions will have the same radial trajectories, but at frequencies dependent on their *m/z* ratio provided that the potential is the same at each frequency [37, 67].

As the excited ions pass near the detector plates, the frequency of their passage is detected as an induced current in the plates; it is called "image" current. This transient signal (known also as "time domain" data or FID, from

"free induction decay") is a composite of the cyclotron frequencies of all the ions present in the cell. The signal slowly dies off with time as the ions relax and return to their stable circular orbits in the center of the analyzer cell, while it is digitized and stored in a computer. The Fourier Transform algorithm is used to extract the frequency and amplitude for each component of the composite signal. The frequency of the motion of an ion is inversely proportional to its mass. This frequency plot is then converted to an *m/z* plot and displayed as the mass spectrum (see Figure 27). The S/N ratio is enhanced by averaging many cycles before transforming and storing data. Once the ions have been detected, a radiofrequency pulse (a quench pulse) is applied to the cell to eject ions before the next bundle of ions are collected and trapped into the cell [39, 76].

FT-ICR-MS is a very powerful tool for analytical proteomics because it can be coupled to an ESI source achieving spectacular resolving power (resolution between two adjacent peaks in a mass spectrum).

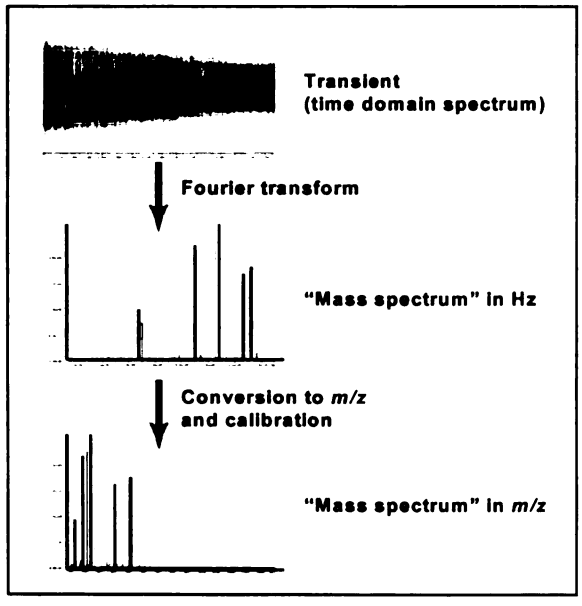


Figure 27. Processing of data from an ICR analyzer using a Fourier transform algorithm, converting time-domain data into frequency domain and then calibrating in terms of *m/z* values (adapted from [109].

3.4 Some examples of instrumentation

3.4a Tandem Mass Spectrometry

Tandem mass spectrometry or MS/MS is a technique in which ion formation/fragmentation and subsequent decomposition of the original ions is carried out "in tandem"; it involves two stages of mass analysis separated by a reaction or fragmentation step.

There are two different approaches to MS/MS experiments:

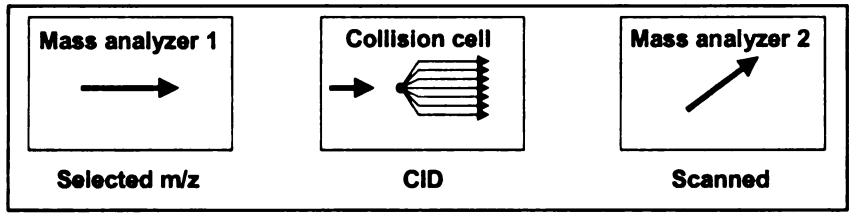
1) Tandem-in-time: Instruments are, in general, ion-trapping mass spectrometers such as two-dimensional or three-dimensional quadrupole ion traps (QIT, LIT) and FT-ICR mass spectrometers. In this case, the various stages of mass spectrometry are conducted within the same physical trapping volume, but at different times during the experiment [98].

2) Tandem-in space: Instruments for these experiments have two (in theory, it could be more than two) mass spectrometers in physically different locations in the instrument.

3.4b Types of MS/MS experiments

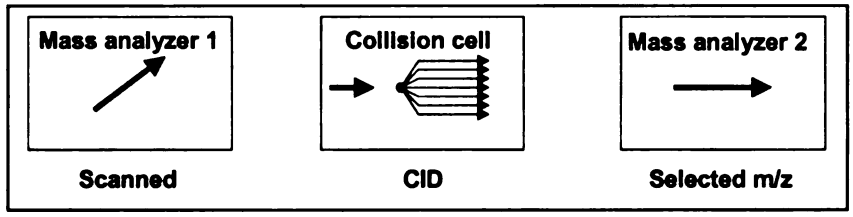
There are four types of MS/MS experiments as described next; conceptual diagrams are also shown for each experiment.

a) Product-ion analysis: This technique involves the isolation of a



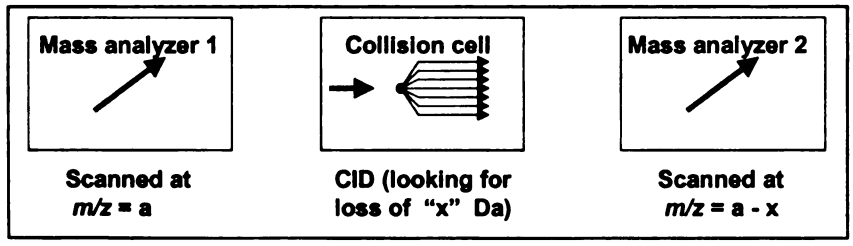
precursor ion, the collisional activated dissociation of the precursor ion and the production of a mass spectrum of the ions resulting from this induced fragmentation.

b) Precursor-ion analysis: All ions formed by a primary ionization pass into



a collision cell, one *m/z* value at a time. The second mass analyzer is set to allow only a single *m/z* value to reach the detector.

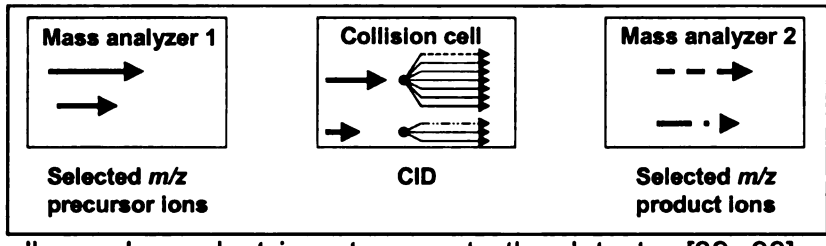
c) Common-neutral-loss analysis: All of the precursor ions formed by a



primary ionization are allowed to pass, one *m/z* value at a time, into

the collision cell. The second mass analyzer scans at the same value of the first mass analyzer, but with a constant *m/z* difference [39].

d) Selected Reaction Monitoring (SRM, sometimes called Multiple Reaction Monitoring, MRM): Ions of specific *m/z* values are allowed to pass to



the collision cell, the second mass analyzer is set to allow only product ions to pass to the detector [39, 98]. This operation mode increases sensitivity and is a good choice for quantitative experiments.

3.4c Tandem-in-time instruments

As it was stated in 3.4a, instruments are, in general, ion-trapping mass spectrometers such as two-dimensional or three-dimensional quadrupole ion traps (QIT, LIT) and FT-ICR mass spectrometers.

The approach for a MS/MS experiment consists of trapping of the ions coming from the ionization source, isolation of the precursor ions of interest by ejecting unwanted ions, fragmenting the precursor ions and performing a final analysis step. In principle, these devices are capable of an unlimited number (n) of repetitive cycles of isolation-fragmentation-isolation of ions before the final

scanning step, thus it might be possible to perform MS/MS/MS....MS experiments (MS^n).

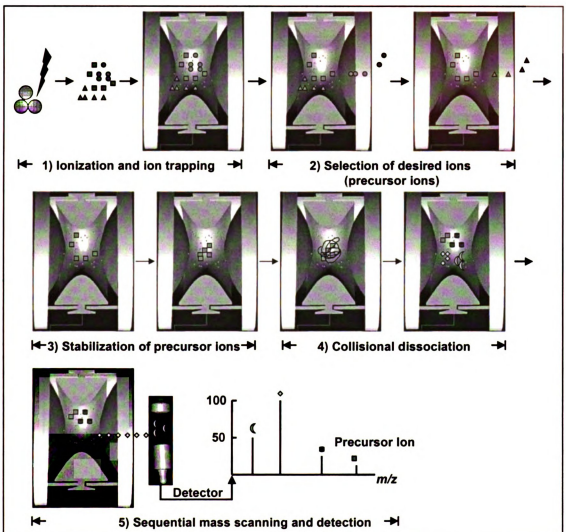


Figure 28. Representation of a MS/MS (MS^2) experiment in a QIT. Idea and QIT drawings from reference [95].

Representation of the MS/MS experiment in Figure 28 can be conceptually extended to a Linear Ion Trap or a FT-ICR.

3.4d Tandem-in-space instruments

Examples of tandem-in-space instruments more commonly used in the analysis of biomolecules are:

Triple quadrupole (QqQ): In the case of the triple quadropole instrument, the collison cell (q) can be a rf-only quadrupole or a hexapole, octapole or a ring guide, however, all these instruments are referred as QqQ.

A very common mode of operation is the Multiple Reaction Monitoring (MRM).

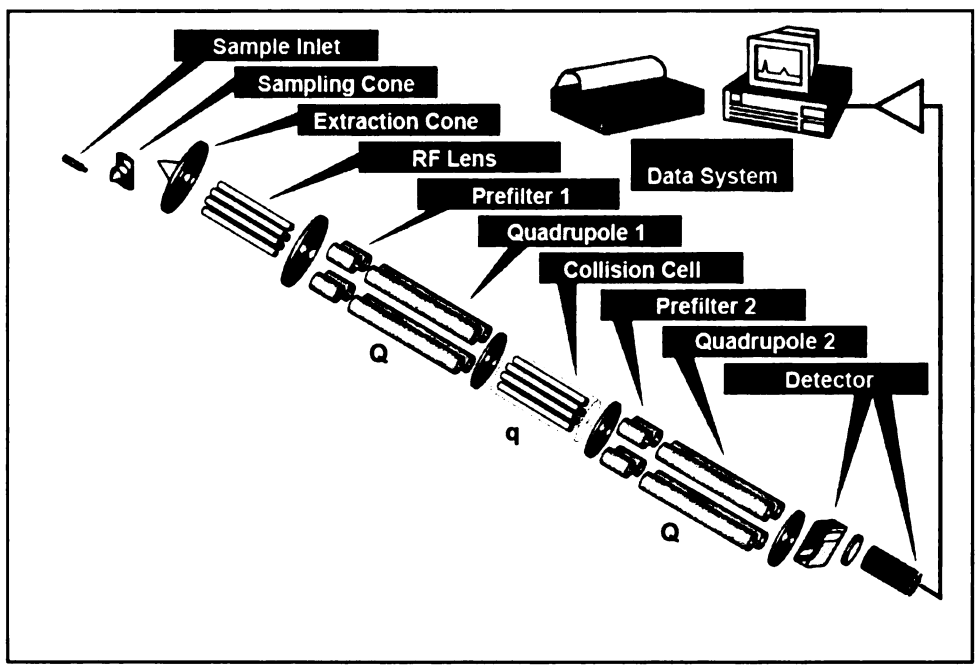


Figure 29. Triple quadrupole scheme for the Quattro LC-MS from Micromass UK Limited [111].

Quadrupole-time-of-flight (QqTOF): This instrument can be thought in terms of replacement of the second mass-selective quadrupole of a QqQ platform with a time-of-flight spectrometer. Using a TOF for the final stage of the mass analysis, the instrument provides high resolution, good mass assignment accuracy, high sensitivity, and the capability to record a complete mass spectrum for each pulse of ions injected into the device [98].

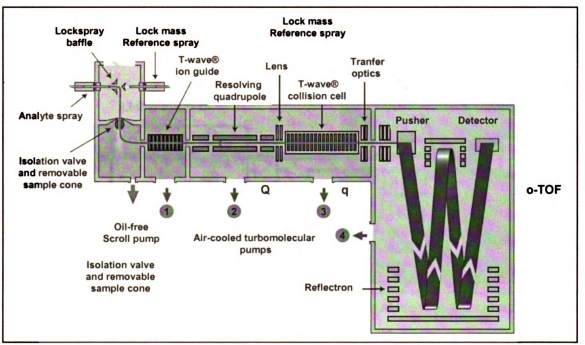


Figure 30. Scheme of the commercial Q-TOF Premier from Micromass/Waters, with an orthogonal TOF analyzer [112].

One variation of a QqTOF instrument is that obtained by substituting the Qq section with a Quadrupole Ion Trap, (QIT-TOF). The use of a linear ion trap instead of a QIT in this assembly has been also proposed (QqLIT); this last setup is interesting because this instrument retains the classical triple quadrupole scan functions such as selected reaction monitoring, product ion, neutral loss, and

precursor ion scanning while also providing access to sensitive ion trap experiments [113].

Linear Traps can be used to improve the performance of FTICR-MS systems. Unwanted ions that can cause space charge problems in the FT-ICR-MS can be ejected from the linear ion trap to improve the resolution, sensitivity, and dynamic range of the system [109].

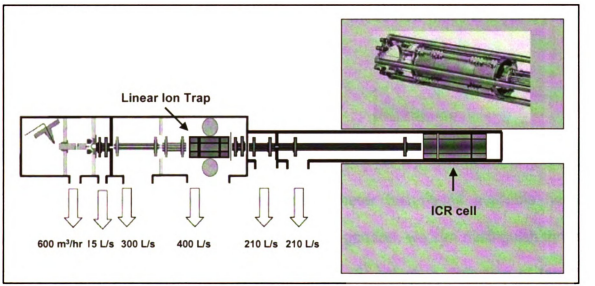


Figure 31. Scheme of the commercial LTQ-FTMS from ThermoFinnigan (It is a LIT-ICR-FT mass spectrometer) [114].

3.4e Other hybrid instruments

Thus far this introduction has described the most common instrumentation that is used in bioanalytical mass spectrometry at this time. However, it is possible to find reports about other possible couplings for MS/MS techniques such as TOF/TOF instruments or others using magnetic (B) and electric sectors (E) in combination with quadrupoles, ion traps or TOF mass spectrometers. In all

cases, the driving force to new designs is the desire to obtain powerful tandem mass instruments with higher resolving power, efficiency, sensitivity, and mass accuracy [67, 98].

During the last few years, new types of mass analyzers have been described, for example, the rectilinear ion trap and the orbitrap invented by Makarov. Recently, a commercial LIT-orbitrap has become available but at this moment it is too early to evaluate the benefits of these new mass analyzers [115-117].

3.4f Applications of MS/MS

The numerous tandem mass spectrometer techniques available have increased the power of mass spectrometry. In general, we can distinguish three types of applications: 1) structure elucidation, 2) selective detection of target compounds, and 3) ion-molecule reactions. In proteomics, genomics, molecular biology, natural products and related fields, MS/MS techniques are becoming more common with the appearance of relatively inexpensive instruments. The understanding of ionization methods, mass analyzers, and detectors is essential in order to take advantage of the capabilities that mass spectrometry offers.

4.- PROTEIN SEQUENCING AND DISULFIDE MASS MAPPING

4.1 Introduction

Even though over the decades, the purpose of peptide sequencing has changed, the main goal has remained intact: the understanding of the function of a protein in a living system. Originally, the aim of peptide sequencing was to determine only the sequence of a protein, wholly or in part, in order to be able to know its structure and by consequence, if possible, its function. As a consequence of the advances in the field of genomics, we approach the time when all genes, and therefore, all protein sequences, will be known. However, knowledge of the sequence of a gene does not give us a complete information on matters that are significant to the function of a protein such as inter- and intramolecular disulfide bonds.

Conventional peptide mapping can be used to determine the entire primary structure of a protein. This method is used routinely to compare the structure of a product to that of a reference material or to those of previous lots to confirm primary structures. In the standard approach, the protein is cleaved at specific sites by enzymes (most notably trypsin) or chemical reagents. The resultant peptides can be easily separated by high performance liquid chromatography (HPLC), two-dimensional thin layer chromatography, or gel electrophoresis. A map or "fingerprint" is obtained, which allows for

differentiation between proteins of similar but not identical, primary structure [118]. If necessary, Edman degradation can be used for the purpose of peptide sequencing of the peptidic fragments. Figure 32 shows an example of peptide mapping using trypsin for the enzymatic cleavage.

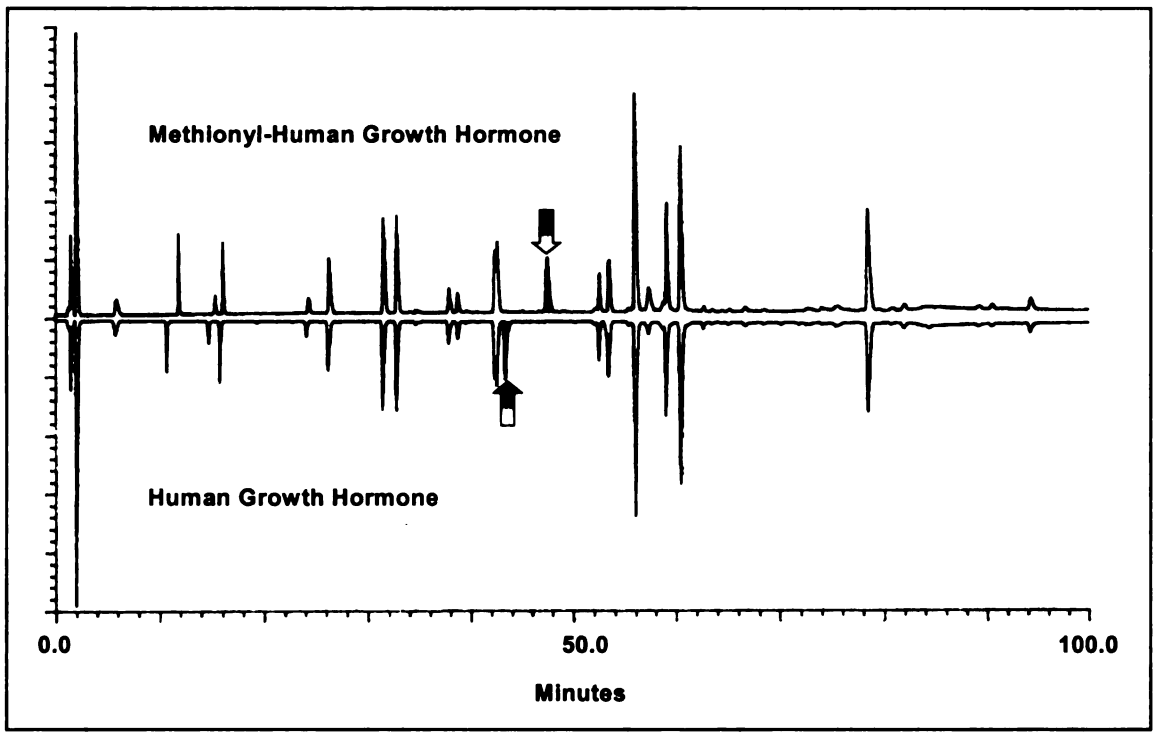


Figure 32. Tryptic map chromatogram of methionyl-human growth hormone and human growth hormone showing the ability of the tryptic map to differentiate the two materials. The N-terminal peptide differences are indicated by the arrows. Slight differences in retention times of peaks eluted early in the chromatogram are attributed to problems with the HPLC-pumps. (From reference [118]).

After the development of Fast Atom Bombardment, it was realized that determination of the molecular weights of peptides generated from proteolytic or chemical digestion of a protein could constitute an identification of the original protein by comparing an experimentally-obtained array of masses with an array of predicted or calculated masses from *in silico* digestion of various sequences listed in a protein database; this technique was called Peptide Mass Mapping

(PMM or peptide mass fingerprinting, PMF). PMM is an extension of the mass mapping technique, the produced peptide fragments, after the protein cleavage, are identified according to their mass. The experimentally obtained masses are compared with the theoretical peptide masses of proteins stored in databases by means of mass search programs [36, 119].

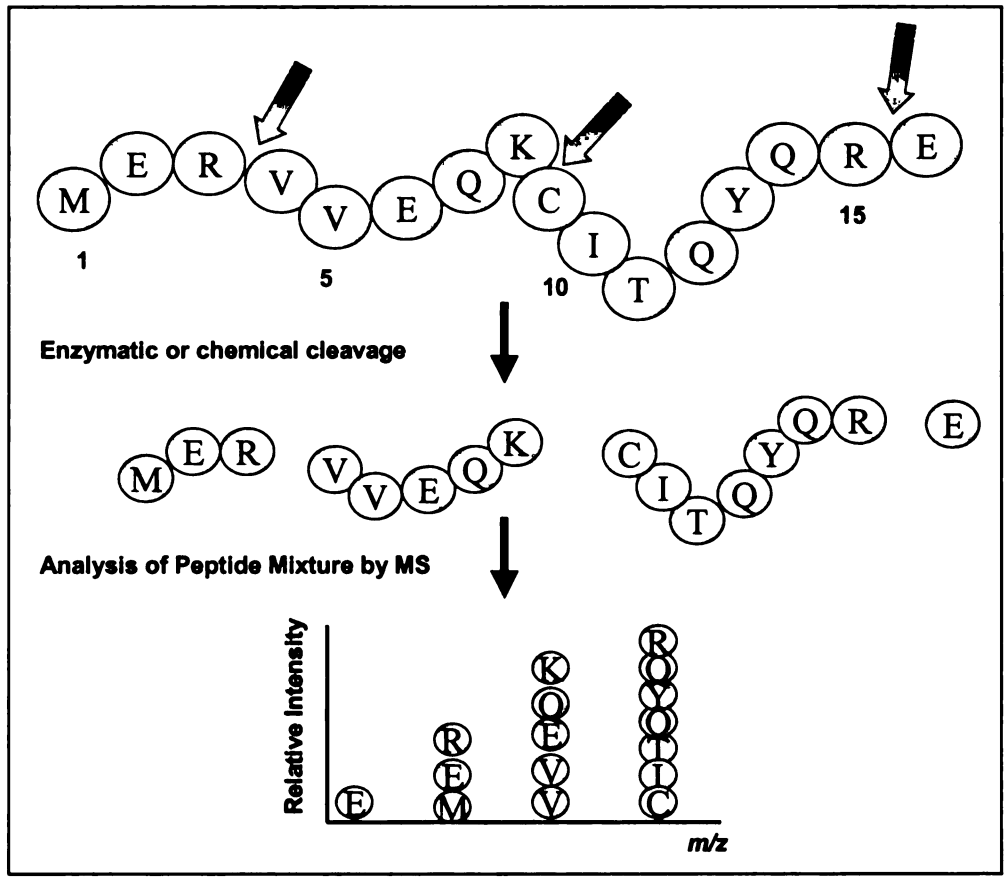


Figure 33. Conceptual diagram of protelytic mass mapping using mass spectrometry. (adapted from reference [36])

Since its description in the late 80s and early 90s, PMM has been used to determine the entire primary structure of proteins, and it has proved to be very useful to provide additional information about posttranslational modifications such as disulfide bonds [118, 120-122].

As we can see from Figure 33, the specific cleavage of a protein yields a specific number of peptides of specific length, sequence, and most importantly of specific mass that are used as inputs for computer algorithms for identification. Currently, there are a number of software internet-based tools available to facilitate protein identification by peptide mass mapping such as Mascot (http://www.matrixscience.com), Profound (http://prowl.rockefeller.edu) ProteinProspector (http://prospector.ucsf.edu), and extensive sequence data bases such as Expasy/SwissProt (http://us.expasy.org/) or NCBI (http://www.ncbi.nlm.nih.gov). MS/MS experiments are also used as a tool for the sequencing and identification of peptide fragments (see Figure 34) [6, 36, 122].

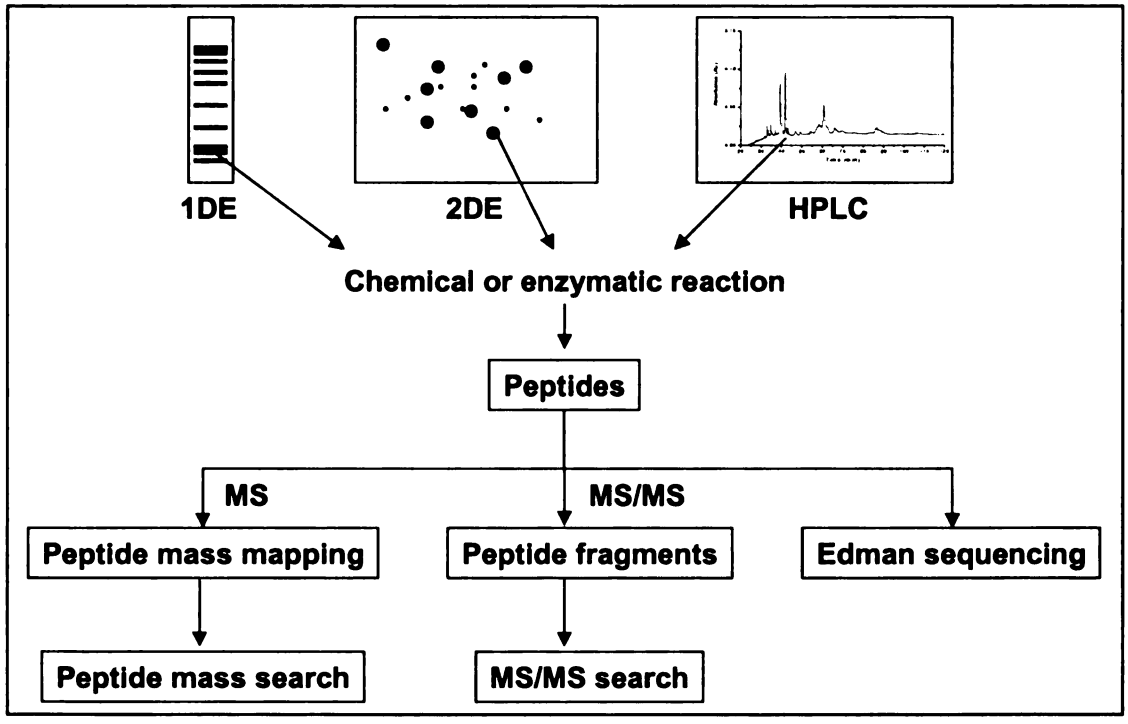


Figure 34. Separation, degradation, and identification techniques used to characterize proteins (adapted from [122]).

Effective PMM is based on several assumptions; the first is that the published genome is accurate. Frame-shifts can arise from incorrect sequencing of the genome, specifically the inclusion of an extra base or the deletion of a base in a gene's sequence. Even with an accurate genome, any of several protein modifications can introduce complications into PMM. Protein or peptide modifications may not be predictable from the genome. These modifications can occur naturally in an organism as a result of deliberate chemical derivatization or via inadvertent modification during sample handling. Another common assumption in peptide mass mapping experiments is that proteolytic or chemical specificity is perfect. Many commonly used proteolytic enzymes and chemical digestion protocols cleave proteins via specific reactions, however, some of these procedures can have side reactions due to impurities (e.g., the presence of chymotripsyn in the trypsin enzyme reagent) or strong chemical conditions (e.g. deamidation at high pH values). Finally, a common assumption in PMM is that all mass spectral features arise from peptides; however, for example, when using MALDI experiments, analyte-matrix or analyte-alkali clusters could produce some confusion during data analysis.

4.2 Disulfide mass mapping

4.2a Importance of disulfide bonds

Analysis of the three-dimensional structure of proteins is important for understanding their biological activity and function. Among the 20 protein amino acids, cysteine has unique properties. It can stabilize the three-dimensional structure of a protein by linking the sulfhydryl groups of two cysteines together to form a disulfide bond. Disulfide linkages can play an essential role in stabilization of protein tertiary structures. Thus, changes in the disulfide bond structure of a protein may result in important changes in activity and function. In addition, the disulfide bond structure of a cystinyl protein may change during protein folding. Knowledge of these structures is a prerequisite to a successful manipulation of complex folding processes, for example, as in biotechnological procedures (e.g., gene transfer) [123].

Disulfide bridges between cysteine residues are a key structural element of many secreted proteins and peptides, being especially abundant in some hormones, enzymes, plasma proteins, inhibitors, and venom proteins. Cysteine can also contribute to protein biological functions because its free sulfhydryl group can play a role in the active site for enzyme catalysis and in chelating metal ions. With some smaller molecules, biological activity may depend strictly on correct pairing of the cysteines. Analyzing the connectivities is thus an

important facet of protein structure determination, but it can take a large investment of time and material. [124, 125]

The advent of Fast Atom Bombardment mass spectrometry during the 1980s made the task of locating disulfide bond more tractable [126]. In the 1990s, new developments in mass spectrometry instrumentation revolutionized protein chemistry; the introduction of electrospray ionization and matrix-assisted laser desorption/ionization [34, 35] solved the problem of generating ions from biomolecules and, therefore, the application of MS in the determination of disulfide bonds has become almost routine.

4.2b Strategies for locating of disulfide bonds in cystinyl proteins

In principle, we can distinguish three different approaches that are currently used for locating of disulfide bonds: enzymatic-based disulfide mass mapping, partial reduction/alkylation-sequencing analysis, and partial reduction/cyanylation CN-induced cleavage mass mapping.

Enzymatic-based disulfide mass mapping

The determination of the disulfide bonds in insulin and ribonuclease-A established the basic strategy that still applies today for the determination of disulfide bonds arrangements using the enzymatic-based approach. The steps

involved in this strategy are: (1) fragmentation of a non-reduced protein of interest into disulfide-linked peptides with cleavages between all half-cystine residues; ideally, the enzyme will cleave the protein at least once between the cysteines as is illustrated in Figure 35, (2) separation of disulfide-linked peptides, (3) location-identification by MS of disulfide-linked peptides, (4) Reduction of disulfide-linked peptides, and (5) isolation-characterization of half-cystinyl peptides using MS [17].

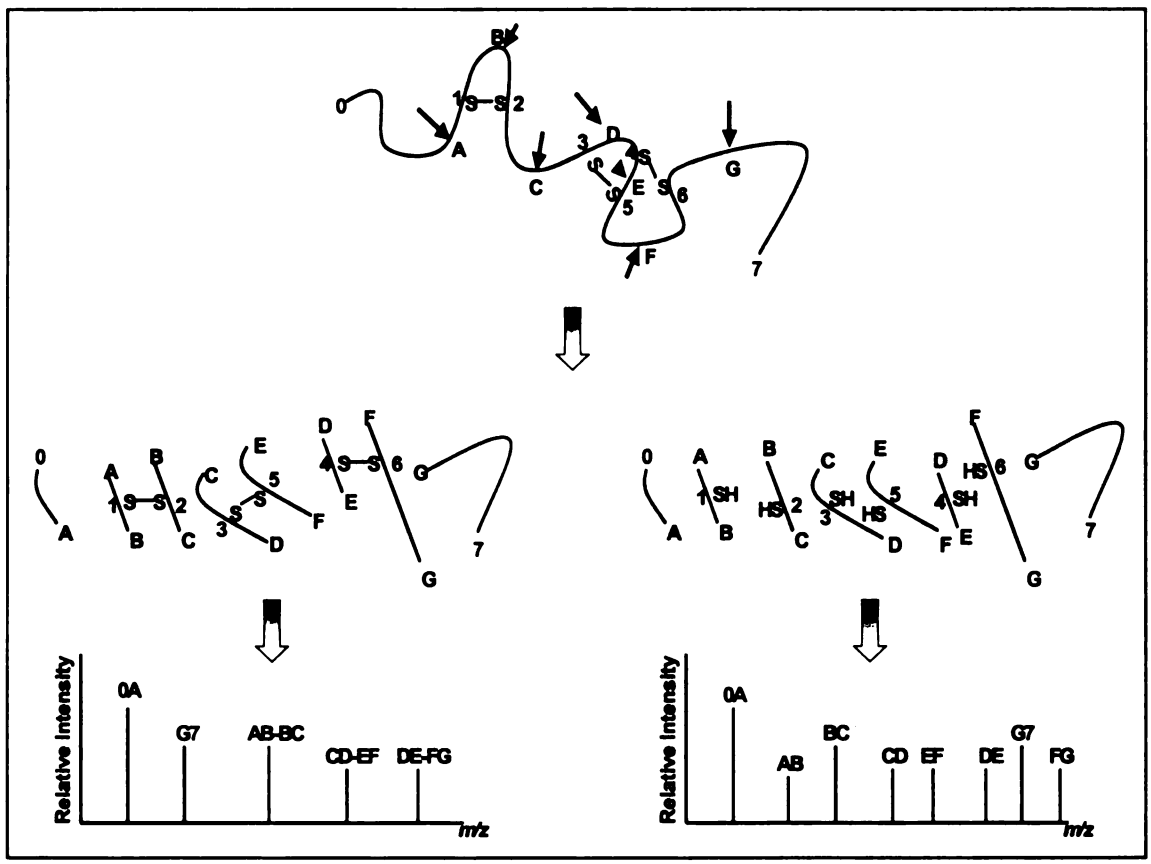


Figure 35. Conceptual scheme illustrating the strategy for determination of disulfide bonds using the enzymatic-based disulfide mass mapping approach.

In this approach, lack of cleavage between each cysteine could complicate the analysis because it will not produce separate proteolytic peptides for each of the cysteines. The closer the cysteines are to one another in the

sequence, the less likely it will be to find a useful proteolytic site; in the case where cysteines are adjacent, this strategy is almost useless. Most proteolytic digestions are carried out in alkaline media causing the possibility of scrambling" or rearrangement (interchange) of native and non-native disulfides [127].

Partial reduction/alkylation-sequencing analysis

The original approach was proposed by Gray [124]. The strategy consists of progressive reduction and alkylation of disulfide bonds followed by characterization by sequence analysis. As the multiple disulfide bonds are reduced by partial reduction, the nascent sulfhydryls are modified with different alkylating reagents to chemically tag the product at different stages of reduction.

This analysis involves 4 key steps: 1) partial reduction in acidic conditions 2) separation of products by HPLC, 3) alkylation of free thiols, and 4) sequence analysis to determine the location of tagged cysteines. Variants of this method were developed in which up to three alkylating agents were used to label different pairs of thiols, allowing in some cases, a full assignment in one sequence analysis (see Figure 36). The incorporation of mass spectrometry has been a straightforward improvement to this methodology. In the original protocol, the use of the water-soluble reducing agent tris-(2 carboxyethyl)-phosphine (TCEP) at pH = 3 was proposed in order to minimize the scrambling problem.

This approach is an advantageous alternative for small cysteine-rich proteins in which only certain disulfides are accessible to the reducing agent. Using this methodology after the reduction of a single disulfide bond, the 3D-protein structure may change and there might be a better accessibility for the reductant during subsequent reduction stages.

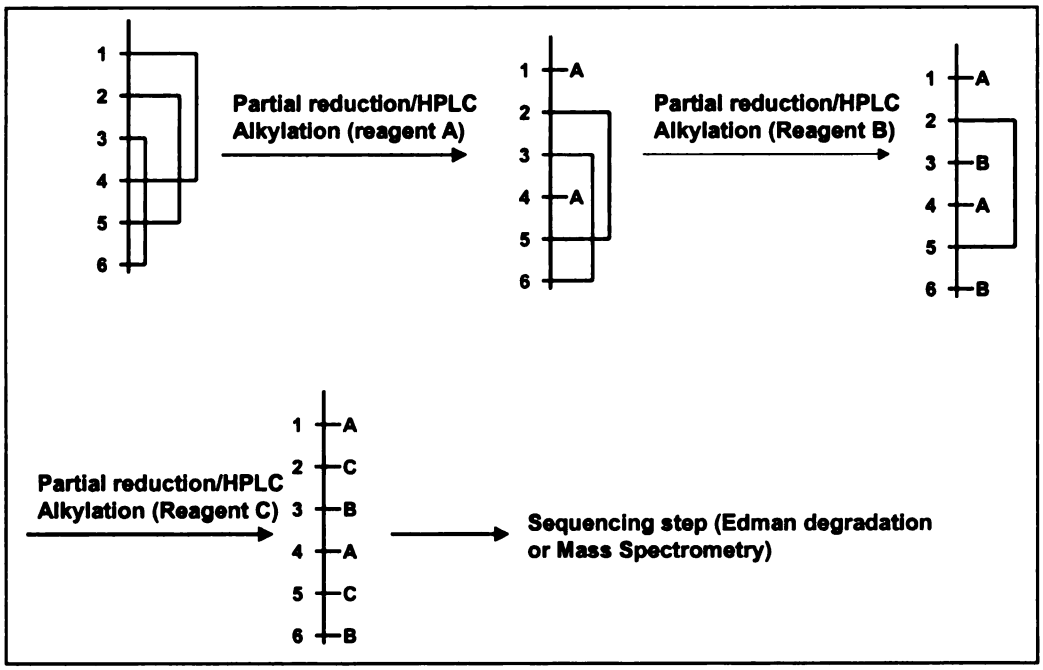


Figure 36. Conceptual scheme of the step-wise partial reduction/alkylation-sequencing method to determine disulfide bond linkage.

This approach has been applied successfully with several modifications according to the specific requirements for a particular protein [128-132]. The most common used alkylation reagents work under slightly alkaline conditions, thus, maybe the most important drawback of this methodology is a possible disulfide exchange during the alkylation step. Nevertheless it seems that the use of maleimides such as N-ethylamaleimide and 4-dimethylaminophenylazophenyl-4'-maleimide) under acidic conditions, can overcome this problem [131-133]

Partial Reduction/Cyanylation CN-induced cleavage (PRCC) based mass mapping methodology

During the late '90s, Wu and Watson established a cyanylation (CN)-based chemical method [125] for disulfide mass mapping to overcome many of the shortcomings of the conventional proteolytic approach. The method is based on five main steps: 1) the cystinyl protein is partially reduced, 2) the nascent cysteines produced during reduction of a given disulfide linkage are cyanylated, 3) the polypeptide backbone is then cleaved on the N-terminal side of the modified cysteines, 4) any residual disulfide bonds are reduced to release the cleavage products, and 5) the reduced cleavage reaction mixture is analyzed by mass spectrometry (MS). Data interpretation leading to the connectivity of cysteines in a disulfide bond relies on recognizing mass spectral peaks that represent CN-induced cleavage products, and mapping these cleavage products to the sequence of the cystinyl protein. During cleavage of the peptide bond on the N-terminal side of a cyano(CN)-modified cysteine, the nucleophile is added to the carbonyl carbon. In this way, the two nascent cysteines in a single-reduced isoform provide the loci of cleavage sites, leading to production of three cleavage fragments: the N-terminal peptide (A), the middle *itz*-peptide (B), and the C-terminal *itz*-peptide (C), in which *itz* represents the 2-iminothiazolidine-4-carboxyl group that is formed during the cleavage step (see Figure 37).

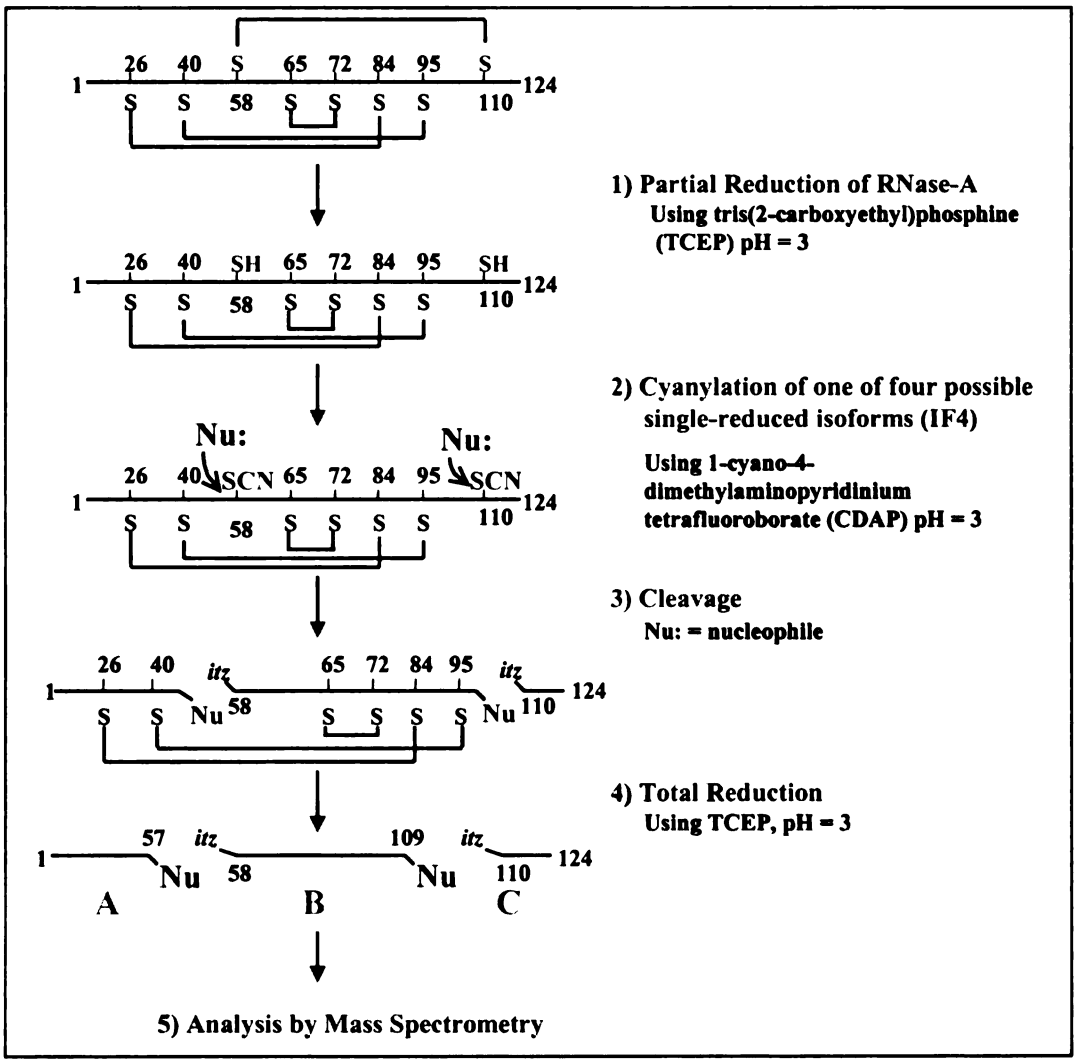


Figure 37. Conceptual representation of the key steps in CN-based methodology for disulfide mass mapping.

The cyanylation (CN)-based mass mapping methodology for assignment of disulfide bond connectivities has proved useful even in proteins containing adjacent cysteines [125, 134, 135], which is one of the greatest challenges in bioanalytical chemistry. Some studies of protein folding are also possible using this methodology [136]. In a step to automate and facilitate data interpretation, an algorithm based on this CN-based methodology has been developed and

applied successfully [137, 138]. In the course of applying the CN-methodology, occasional disappointing yields of certain cleavage products have been observed, the cause of which was apparently related to some complex function of amino acid composition and sequence in the vicinity of the cysteines [139]. The analysis of the cleavage reaction is the subject of chapter 2.

Miscellaneous methods

There are other methods that have been proposed for determining disulfide connectivity; in general, they are only variations of the methods described above with modifications to minimize the risk of protein scrambling or with the goal of improving yields and detection of proteolytic or chemical-induced fragments. Maybe the only novel and promising approach for determination of disulfide linkage is by using electron capture dissociation [78, 140]; however, at this time it is too early to know the impact of this technique in the determination of disulfide bonds in cystinyl proteins.

5.- SUMMARY

Undoubtedly modern mass spectrometry combined with chemical derivatization is a very powerful tool in the analysis of biomolecules and currently, it is the most important analytical tool used in proteomics. In this dissertation, the use of mass spectrometry in cyanylation-based disulfide mass mapping give us an excellent example of the application of mass spectrometry in biomedical science and in the determination of the disulfide proteome. This methodology relies on the chemical derivatization and cleavage of the protein prior analysis my mass spectrometry.

In Chapter 2 of this dissertation is presented a systematic study of the nucleophile-mediated CN-induced cleavage reaction for the cyanylation-based methodology using primary (methyl to butyl) amines as an alternative to ammonia as the cleavage reagent. Reaction conditions are systematically studied using high performance liquid chromatography to monitor the yield of the expected cleavage products as identified by mass spectrometry. The focus of this study was to optimize cleavage of cyanylated cystinyl proteins and to improve the detection of the cleavage products.

In Chapter 3 is investigated the use of binary mixtures of homologous nucleophiles to improve the confidence in distinguishing CN-induced cleavage products from chemical by-products during analyses of cleavage reaction

mixtures by MALDI-MS. Analytical advantage can be gained in the use of a binary mixture of homologous nucleophiles for cyanylation-induced cleavage of partially reduced isoforms of cystinyl proteins. The combined use of ammonia and methylamine was studied in order to find the adequate concentration ratio for labeling of cleavage products in a single chemical step. The appearance of pairs of mass spectral peaks facilitates recognition of diagnostic CN-induced cleavage products, and increases confidence in results when samples present weak signals or interfering chemical noise.

Chapter 4 presents a preliminary development of an on-line system for the cyanylation-based methodology. The goal of this development was to carry out all the chemistry on-line in order to avoid sample manipulation and, by consequence, possible sample losses. The on-line system consists of a first trap filled with a polymeric resin in which the protein is bound, partially reduced and cyanylated. After these steps, the protein is eluted from the first trap and transferred into a reactor for the cleavage reaction. Finally, the protein cleavage mixture is trapped onto a second polymeric column where the residual disulfide bonds are reduced and the remaining salts are removed from the sample. The second trap containing the CN-induced cleavage fragments can be connected to a HPLC or a LC-MS system. Cleavage fragments also can be eluted to be analyzed by MALDI-TOF-MS.

In Chapter 5 of this dissertation are presented conclusions and future work is sketched.

6.- REFERENCES

[1] T. E. Creighton, Proteins. Structure and Molecular Principles, W. H. Freeman, New York, NY., 1993.

[2] M. Trudy, M. R. James, Biochemistry, McGraw-Hill, Beijing, People's Republic of China., 1999.

[3] C. Cambillau, Proteins - reflecting the diversity of life, *Curr. Opin. Struct. Biol.* 13 (2003) 655-657.

[4] A. L. Fink, Natively unfolded proteins, *Curr. Opin. Struct. Biol.* 15 (2005) 35-41.

[5] V. N. Uversky, Natively unfolded proteins: a point where biology waits for physics, *Protein Sci.* 11 (2002) 739-756.

[6] D. C. Liebler, Editor, Introduction to Proteomics: Tools for the New Biology, Humana Press, Totowa, NJ., 2002.

[7] M. R. Wilkins, J. C. Sanchez, A. A. Gooley, R. D. Appel, I. Humphery-Smith, D. F. Hochstrasser, K. L. Williams, Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it, *Biotechnol. Genet. Eng. Rev.* 13 (1996) 19-50.

[8] J. S. Andersen, M. Mann, Functional genomics by mass spectrometry, *FEBS Lett.* 480 (2000) 25-31.

[9] P. R. Graves, T. A. Haystead, Molecular biologist's guide to proteomics, *Microbiol. Mol. Biol. Rev.* 66 (2002) 39-63;.

[10] K. Sperling, From proteomics to genomics, *Electrophoresis* 22 (2001) 2835-2837.

[11] M. Wang, R. J. Lamers, H. A. Korthout, J. H. van Nesselrooij, R. F. Witkamp, R. van der Heijden, P. J. Voshol, L. M. Havekes, R. Verpoorte, J. van der Greef, Metabolomics in the context of systems biology: bridging traditional Chinese medicine and molecular pharmacology, *Phytother. Res.* 19 (2005) 173-182.

[12] A. Pandey, M. Mann, Proteomics to study genes and genomes, *Nature* 405 (2000) 837-846.

[13] H. Yano, S. Kuroda, B. B. Buchanan, Disulfide proteome in the analysis of protein function and structure, *Proteomics* 2 (2002) 1090-1096.

[14] A. Hiniker, J. C. Bardwell, Disulfide bond isomerization in prokaryotes, *Biochemistry* 42 (2003) 1179-1185.

[15] R. L. Baldwin, Protein folding from 1961 to 1982, *Nat. Struct. Biol.* 6 (1999) 814-817.

[16] E. Welker, M. Narayan, W. J. Wedemeyer, H. A. Scheraga, Structural determinants of oxidative folding in proteins, *Proc. Natl. Acad. Sci. U.S.A.* 98 (2001) 2312-2316.

[17] J. J. Gorman, T. P. Wallis, J. J. Pitt, Protein disulfide bond determination by mass spectrometry, *Mass Spectrom. Rev.* 21 (2002) 183-216.

[18] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags, *Nat. Biotechnol.* 17 (1999) 994-999.

[19] P. L. Martelli, P. Fariselli, R. Casadio, Prediction of disulfide-bonded cysteines in proteomes with a hidden neural network, *Proteomics* 4 (2004) 1665-1671.

[20] J. P. Bingham, N. M. Broxton, B. G. Livett, J. G. Down, A. Jones, E. G. Moczydlowski, Optimizing the connectivity in disulfide-rich peptides: alpha-conotoxin SII as a case study, *Anal. Biochem.* 338 (2005) 48-61.

[21] E. Fuller, B. R. Green, P. Catlin, O. Buczek, J. S. Nielsen, B. M. Olivera, G. Bulaj, Oxidative folding of conotoxins sharing an identical disulfide bridging framework, *FEBS J.* 272 (2005) 1727-1738.

[22] X. Chen, Y. Li, X. Tong, N. Zhang, G. Wu, Q. Zhang, H. Wu, Solution structure of BmP08, a novel short-chain scorpion toxin from Buthus martensi Karsch, *Biochem. Biophys. Res. Commun.* 330 (2005) 1116-1126.

[23] J. J. Calvete, C. Marcinkiewicz, D. Monleon, V. Esteve, B. Celda, P. Juarez, L. Sanz, Snake venom disintegrins: evolution of structure and function, *Toxicon* 45 (2005) 1063-1074.

[24] Y. Xiao, S. Liang, Inhibition of neuronal tetrodotoxin-sensitive Na+ channels by two spider toxins: hainantoxin-III and hainantoxin-IV, *Eur. J. Pharmacol.* 477 (2003) 1-7.

[25] E. Welker, W. J. Wedemeyer, H. A. Scheraga, A role for intermolecular disulfide bonds in prion diseases?, *Proc. Natl. Acad. Sci. U.S.A.* 98 (2001) 4334-4336.

[26] T. P. Wallis, J. J. Pitt, J. J. Gorman, Identification of disulfide-linked peptides by isotope profiles produced by peptic digestion of proteins in 50% (18)O water, *Protein Sci.* 10 (2001) 2251-2271.

[27] J. J. Pitt, E. Da Silva, J. J. Gorman, Determination of the disulfide bond arrangement of Newcastle disease virus hemagglutinin neuraminidase. Correlation with a beta-sheet propeller structural fold predicted for paramyxoviridae attachment proteins, *J. Biol. Chem.* 275 (2000) 6469-6478.

[28] P. Eaton, M. E. Jones, E. McGregor, M. J. Dunn, N. Leeds, H. L. Byers, K. Y. Leung, M. A. Ward, J. R. Pratt, M. J. Shattock, Reversible cysteine-targeted oxidation of proteins during renal oxidative stress, *J. Am. Soc. Nephrol.* 14 (2003) S290-296.

[29] H. P. Harding, Y. Zhang, H. Zeng, I. Novoa, P. D. Lu, M. Calfon, N. Sadri, C. Yun, B. Popko, R. Paules, D. F. Stojdl, J. C. Bell, T. Hettmann, J. M. Leiden, D. Ron, An integrated stress response regulates amino acid metabolism and resistance to oxidative stress, *Mol. Cell* 11 (2003) 619-633.

[30] M. Xu, A. Arulandu, D. K. Struck, S. Swanson, J. C. Sacchettini, R. Young, Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme, *Science* 307 (2005) 113-117.

[31] K. Lee, J. Lee, Y. Kim, D. Bae, K. Y. Kang, S. C. Yoon, D. Lim, Defining the plant disulfide proteome, *Electrophoresis* 25 (2004) 532-541.

[32] E. de Pauw, Matrix selection for liquid secondary ion and fast atom bombardment mass spectrometry, *Methods Enzymol.* 193 (1990) 201-214.

[33] J. Schiller, K. Arnold (2000) *in* Encyclopedia of Analytical Chemistry (Meyers, R. A., Ed.), pp. 559-585, John Wiley & Sons Ltd, Chichester, NY.

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

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

[36] J. T. Watson, Introduction to Mass Spectrometry, Lippincott Williams & Wilkins, Philadelphia, PA., 1997.

[37] IonSpec.com. FTMS tutorial part I, http://www.ionspec.com/FTMS%20Tutorial/tutorial_fs.htm, (accessed August 15 2005).

[38] A. S. Carr, S. R. Annan, Overview of Peptide and Protein Analysis by Mass Spectrometry, John Wiley @ Sons, Inc., 1997.

[39] D. Sparkman O., Mass Spectrometry Desk Reference, Global View Publishing, Pittsburgh, PA., 2000.

[40] R. A. Zubarev, Reactions of polypeptide ions with electrons in the gas phase, *Mass Spectrom. Rev.* 22 (2003) 57-77.

[41] W. E. Seifert, Jr., R. M. Caprioli, Fast atom bombardment mass spectrometry, *Methods Enzymol.* 270 (1996) 453-486.

[42] F. Hillenkamp, U. Bahr, M. Karas, B. Spengler, Mechanisms of laser ion formation for mass spectrometric analysis, *Scanning Microscopy, Supplement* 1 (1987) 33-39.

[43] R. Holm, M. Karas, H. Vogt, Polymer investigations with the laser microprobe, *Anal. Chem.* 59 (1987) 371-373.

[44] M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Matrix-assisted ultraviolet laser desorption of non-volatile compounds, *Int. J. Mass Spectrom. Ion Processes* 78 (1987) 53-68.

[45] Sigma-Aldrich (2001) MALDI-Mass Spectrometry *in* Analytix.

[46] A. Schmidt, M. Karas, T. Dulcks, Effect of different solution flow rates on analyte ion signals in nano-ESI MS, or: when does ESI turn into nano-ESI?, *J. Am. Soc. Mass Spectrom.* 14 (2003) 492-500.

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

[48] R. Zenobi, R. Knochenmuss, Ion formation in MALDI mass spectrometry, *Mass Spectrom. Rev.* 17 (1998) 337-366.

[49] R. Knochenmuss, A quantitative model of ultraviolet matrix-assisted laser desorption/ionization, *J. Mass Spectrom.* 37 (2002) 867-877.

[50] R. Knochenmuss, A quantitative model of ultraviolet matrix-assisted laser desorption/ionization including analyte ion generation, *Anal. Chem.* 75 (2003) 2199-2207.

[51] M. Cadene, B. T. Chait, A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins, *Anal. Chem.* 72 (2000) 5655-5658.

[52] J. J. Thomas, Z. Shen, J. E. Crowell, M. G. Finn, G. Siuzdak, Desorption/ionization on silicon (DIOS): a diverse mass spectrometry platform for protein characterization, *Proc. Natl. Acad. Sci. U.S.A.* 98 (2001) 4932-4937.

[53] J. Wei, J. M. Buriak, G. Siuzdak, Desorption-ionization mass spectrometry on porous silicon, *Nature* 399 (1999) 243-246.

[54] Y. Xu, M. L. Bruening, J. T. Watson, Non-specific, on-probe cleanup methods for MALDI-MS samples, *Mass Spectrom. Rev.* 22 (2003) 429-440.

[55] Y. Xu, J. T. Watson, M. L. Bruening, Patterned monolayer/polymer films for analysis of dilute or salt-contaminated protein samples by MALDI-MS, *Anal. Chem.* 75 (2003) 185-190.

[56] Y. Xu, M. L. Bruening, J. T. Watson, Use of polymer-modified MALDI-MS probes to improve analyses of protein digests and DNA, *Anal. Chem.* 76 (2004) 3106-3111.

[57] M. Dole, R. L. Hines, L. L. Mack, R. C. Mobley, L. D. Ferguson, M. B. Alice, Gas phase macroions, *Macromolecules* 1 (1968) 96-97.

[58] M. Dole, L. L. Mack, R. L. Hines, R. C. Mobley, L. D. Ferguson, M. B. Alice, Molecular beams of macroions, *J. Chem. Phys.* 49 (1968) 2240-2249.

[59] M. Yamashita, J. B. Fenn, Electrospray ion source. Another variation on the free-jet theme, *J. Phys. Chem.* 88 (1984) 4451-4459.

[60] M. Yamashita, J. B. Fenn, Negative ion production with the electrospray ion source, *J. Phys. Chem.* 88 (1984) 4671-4675.

[61] M. Yamashita, J. B. Fenn, Application of electrospray mass spectrometry in medicine and biochemistry, *Koenchu-Iyo Masu Kenkyukai* 9 (1984) 203-206.

[62] C. K. Meng, M. Mann, J. B. Fenn, Of protons or proteins, *Zeitschrift fuer Physik D: Atoms, Molecules and Clusters* 10 (1988) 361-368.

[63] M. L. Aleksandrov, L. N. Gall, N. V. Krasnov, V. I. Nikolaev, V. A. Pavlenko, V. A. Shkurov, G. I. Baram, M. A. Grachev, V. D. Knorre, Y. S. Kusner, Direct coupling of a microcolumn liquid chromatograph and a mass spectrometer, *Bioorg. Khim.* 10 (1984) 710-712.

[64] M. Wilm, M. Mann, Analytical properties of the nanoelectrospray ion source, *Anal. Chem.* 68 (1996) 1-8.

[65] E. d. Hoffmann, J. J. Charette, V. Stroobant, Spectrométrie de masse, Dunod Ed., Paris, 1999.

[66] J. H. Gross, Mass spectrometry: a textbook, Springer, Berlin ; New York, 2004.

[67] E. d. Hoffmann, V. Stroobant, Mass spectrometry : principles and applications, Wiley, Chichester ; New York, 2001.

[68] J. B. Fenn, Electrospray wings for molecular elephants (Nobel lecture), *Angew. Chem. Int. Ed. Engl.* 42 (2003) 3871-3894.

[69] L. Tang, P. Kebarle, Dependence of ion intensity in electrospray mass spectrometry on the concentration of the analytes in the electrosprayed solution, *Anal. Chem.* 65 (1993) 3654-3668.

[70] P. Kebarle, L. Tang, From ions in solution to ions in the gas phase - the mechanism of electrospray mass spectrometry, *Anal. Chem.* 65 (1993) 972A-986A.

[71] P. Kebarle, A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry, *J. Mass Spectrom.* 35 (2000) 804-817.

[72] P. Kebarle, M. Peschke, On the mechanisms by which the charged droplets produced by electrospray lead to gas phase ions, *Anal. Chim. Acta* 406 (2000) 11-35.

[73] J. V. Iribarne, B. A. Thomson, On the evaporation of small ions from charged droplets, *J. Chem. Phys.* 64 (1976) 2287-2294.

[74] M. R. Emmett, R. M. Caprioli, Micro-electrospray mass spectrometry: Ultra-high-sensitivity analysis of peptides and proteins, *J. Am. Soc. Mass Spectrom.* 5 (1994) 605-613.

[75] M. S. Wilm, M. Mann, Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last?, *Int. J. Mass Spectrom. Ion Processes* 136 (1994) 167-180.

[76] R. C. Willoughby, E. Sheehan, S. Mitrovich, A global view of LC/MS: how to solve your most challenging analytical problems, Global View Publ., Pittsburgh, Pa., 2002.

[77] R. A. Zubarev, N. L. Kelleher, F. W. McLafferty, Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process, *J. Am. Chem. Soc.* 120 (1998) 3265-3266.

[78] R. A. Zubarev, Electron-capture dissociation tandem mass spectrometry, *Curr. Opin. Biotechnol.* 15 (2004) 12-16.

[79] F. Turecek, N-Ca Bond Dissociation Energies and Kinetics in Amide and Peptide Radicals. Is the Dissociation a Non-ergodic Process?, *J. Am. Chem. Soc.* 125 (2003) 5954-5963.

[80] E. A. Syrstad, F. Turecek, Toward a general mechanism of electron capture dissociation, *J. Am. Soc. Mass Spectrom.* 16 (2005) 208-224.

[81] D. M. Horn, R. A. Zubarev, F. W. McLafferty, Automated de novo sequencing of proteins by tandem high-resolution mass spectrometry, *Proc. Natl. Acad. Sci. U.S.A.* 97 (2000) 10313-10317.

[82] K. Breuker, H. Oh, C. Lin, B. K. Carpenter, F. W. McLafferty, Nonergodic and conformational control of the electron capture dissociation of protein cations, *Proc. Natl. Acad. Sci. U.S.A.* 101 (2004) 14011-14016.

[83] H. J. Cooper, K. Hakansson, A. G. Marshall, The role of electron capture dissociation in biomolecular analysis, *Mass Spectrom. Rev.* 24 (2005) 201-222.

[84] D. M. Horn, Y. Ge, F. W. McLafferty, Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins, *Anal. Chem.* 72 (2000) 4778-4784.

[85] S. K. Sze, Y. Ge, H. Oh, F. W. McLafferty, Top-down mass spectrometry of a 29-kDa protein for characterization of any posttranslational modification to within one residue, *Proc. Natl. Acad. Sci. U.S.A.* 99 (2002) 1774-1779.

[86] Y. Ge, B. G. Lawhorn, M. ElNaggar, E. Strauss, J. H. Park, T. P. Begley, F. W. McLafferty, Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry, *J. Am. Chem. Soc.* 124 (2002) 672-678.

[87] R. A. Zubarev, N. A. Kruger, E. K. Fridriksson, M. A. Lewis, D. M. Horn, B. K. Carpenter, F. W. McLafferty, Electron Capture Dissociation of Gaseous Multiply-Charged Proteins Is Favored at Disulfide Bonds and Other Sites of High Hydrogen Atom Affinity, *Journal of the American Chemical Society* 121 (1999) 2857-2862.

[88] K. R. Jennings, G. G. Dolnikowski, Mass analyzers, *Methods Enzymol.* 193 (1990) 37-61.

[89] R. M. Smith, K. L. Busch, Understanding mass spectra: a basic approach, Wiley, New York, NY., 1999.

[90] C. Steel, M. Henchman, Understanding the quadrupole mass filter through computer simulation, *J. Chem. Educ.* 75 (1998) 1049-1054.

[91] J. J. Leary, R. L. Schmidt, Quadrupole mass spectrometers: an intuitive look at the math, *J. Chem. Educ.* 73 (1996) 1142-1145.

[92] P. E. Miller, M. B. Denton, The quadrupole mass filter: basic operating concepts, *J. Chem. Educ.* 63 (1986) 617-622.

[93] J. Pappas *Quadrupole Ion-Trap Mass Analyzer*, Thermo Electron, Somerset, NJ., (2004).

[94] M. W. Forbes, M. Sharifi, T. Croley, Z. Lausevic, R. E. March, Simulation of ion trajectories in a quadrupole ion trap: a comparison of three simulation programs, *J. Mass Spectrom.* 34 (1999) 1219-1239.

[95] Ion trap demo, Thermo-Finnigan, 2002.

[96] D. J. Douglas, A. J. Frank, D. Mao, Linear ion traps in mass spectrometry, *Mass Spectrom. Rev.* 24 (2005) 1-29.

[97] J. W. Hager, A new linear ion trap mass spectrometer, *Rapid Commun. Mass Spectrom.* 16 (2002) 512-526.

[98] J. W. Hager, Recent trends in mass spectrometer development, *Anal. Bioanal. Chem.* 378 (2004) 845-850.

[99] J. C. Schwartz, M. W. Senko, J. E. Syka, A two-dimensional quadrupole ion trap mass spectrometer, *J. Am. Soc. Mass Spectrom.* 13 (2002) 659-669.

[100] M. Sudakov, D. J. Douglas, Linear quadrupoles with added octopole fields, *Rapid Commun. Mass Spectrom.* 17 (2003) 2290-2294.

[101] V. Mayya, K. Rezaul, Y. S. Cong, D. Han, Systematic comparison of a two-dimensional ion trap and a three-dimensional ion trap mass spectrometer in proteomics, *Mol. Cell. Proteomics* 4 (2005) 214-223.

[102] W. C. Wiley, I. H. McLaren, Time-of-flight mass spectrometer with improved resolution, *Rev. Sci. Instr.* 26 (1955) 1150-1157.

[103] M. Karas, \"Time-of-flight mass spectrometer with improved resolution,\" W. C. Wiley and I. H. McLaren, Rev. Sci. Instrum., 26, 1150 (1955), *J. Mass Spectrom.* 32 (1997) 1-3.

[104] M. Guilhaus, V. Mlynski, D. Selby, Perfect timing: time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.* 11 (1997) 951-962.

[105] D. C. Muddiman, R. Bakhtiar, S. A. Hofstadler, R. D. Smith, Matrix-assisted laser desorption/ionization mass spectrometry: instrumentation and applications, *J. Chem. Educ.* 74 (1997) 1288-1292.

[106] R. J. Cotter, The new time-of-flight mass spectrometry, *Anal. Chem.* 71 (1999) 445A-451A.

[107] B. A. Mamyrin, Laser assisted reflectron time-of-flight mass spectrometry, *Int. J. Mass Spectrom. Ion Processes* 131 (1994) 1-19.

[108] J. H. J. Dawson, M. Guilhaus, Orthogonal-acceleration time-of-flight mass spectrometer, *Rapid Commun. Mass Spectrom.* 3 (1989) 155-159.

[109] M. P. Barrow, W. I. Burkitt, P. J. Derrick, Principles of Fourier transform ion cyclotron resonance mass spectrometry and its application in structural biology, *Analyst* 130 (2005) 18-28.

[110] A. G. Marshall, C. L. Hendrickson, G. S. Jackson, Fourier transform ion cyclotron resonance mass spectrometry: a primer, *Mass Spectrom. Rev.* 17 (1998) 1-35.

[111] MicroMass UK Limited. Quattro LC User's Guide, http://www.micromass.co.uk, (accessed August 2005).

[112] W. MicroMass UK Limited. Q-Tof premier mass spectrometer (brochure), http://www.waters.com, (accessed August 2005).

[113] G. Hopfgartner, E. Varesio, V. Tschappat, C. Grivet, E. Bourgogne, L. A. Leuthold, Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules, *J. Mass Spectrom.* 39 (2004) 845-855.

[114] Thermo-Finnigan. LTQ-FTMS, (accessed August 2005).

[115] Q. Hu, R. J. Noll, H. Li, A. Makarov, M. Hardman, R. Graham Cooks, The Orbitrap: a new mass spectrometer, *J. Mass Spectrom.* 40 (2005) 430-443.

[116] M. Hardman, A. A. Makarov, Interfacing the orbitrap mass analyzer to an electrospray ion source, *Anal. Chem.* 75 (2003) 1699-1705.

[117] Z. Ouyang, G. Wu, Y. Song, H. Li, W. R. Plass, R. G. Cooks, Rectilinear ion trap: concepts, calculations, and analytical performance of a new mass analyzer, *Anal. Chem.* 76 (2004) 4595-4605.

[118] R. L. Garnick, N. J. Solli, P. A. Papa, The role of quality control in biotechnology: an analytical perspective, *Anal. Chem.* 60 (1988) 2546-2557.

[119] Bioinformatics Web. Genomic glossary, http://www.geocities.com/bioinformaticsweb/index.html, (accessed August 2005).

[120] W. J. Henzel, T. M. Billeci, J. T. Stults, S. C. Wong, C. Grimley, C. Watanabe, Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases, *Proc. Natl. Acad. Sci. U.S.A.* 90 (1993) 5011-5015.

[121] B. W. Gibson, K. Biemann, Strategy for the mass spectrometric verification and correction of the primary structures of proteins deduced from their DNA sequences, *Proc. Natl. Acad. Sci. U.S.A.* 81 (1984) 1956-1960.

[122] B. Thiede, W. Hohenwarter, A. Krah, J. Mattow, M. Schmid, F. Schmidt, P. R. Jungblut, Peptide mass fingerprinting, *Methods* 35 (2005) 237-247.

[123] S. Barbirz, H. P. Happersberger, M. Przybylski, M. O. Glocker, Selective cyanylation of cysteinyl residues as an approach for the mass spectrometric determination of protein structures, *European Mass Spectrometry* 5 (1999) 123-131.

[124] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, How to measure and predict the molar absorption coefficient of a protein, *Protein Sci.* 4 (1995) 2411-2423.

[125] J. Wu, J. T. Watson, A novel methodology for assignment of disulfide bond pairings in proteins, *Protein Sci.* 6 (1997) 391-398.

[126] D. L. Smith, Z. R. Zhou, Strategies for locating disulfide bonds in proteins, *Methods Enzymol.* 193 (1990) 374-389.

[127] J. Wu, J. Qi, J. T. Watson, Encyclopedia of mass spectrometry, in press.

[128] J. E. Hale, J. P. Butler, V. Gelfanova, J. S. You, M. D. Knierman, A simplified procedure for the reduction and alkylation of cysteine residues in proteins prior to proteolytic digestion and mass spectral analysis, *Anal. Biochem.* 333 (2004) 174-181.

[129] V. Schnaible, S. Wefing, A. Bucker, S. Wolf-Kummeth, D. Hoffmann, Partial reduction and two-step modification of proteins for identification of disulfide bonds, *Anal. Chem.* 74 (2002) 2386-2393.

[130] S. Sechi, B. T. Chait, Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification, *Anal. Chem.* 70 (1998) 5150-5158.

[131] U. Goransson, A. M. Broussalis, P. Claeson, Expression of Viola cyclotides by liquid chromatography-mass spectrometry and tandem mass spectrometry sequencing of intercysteine loops after introduction of charges and cleavage sites by aminoethylation, *Anal. Biochem.* 318 (2003) 107-117.

[132] U. Goransson, D. J. Craik, Disulfide mapping of the cyclotide kalata B1. Chemical proof of the cystic cystine knot motif, *J. Biol. Chem.* 278 (2003) 48188-48196.

[133] C. R. Borges, J. T. Watson, Recognition of cysteine-containing peptides through prompt fragmentation of the 4-dimethylaminophenylazophenyl-4'-maleimide derivative during analysis by MALDI-MS, *Protein Sci.* 12 (2003) 1567-1572.

[134] J. Wu, J. T. Watson, Assignment of disulfide bonds in proteins by chemical cleavage and peptide mapping by mass spectrometry, *Methods in Molecular Biology (Totowa, NJ, United States)* 194 (2002) 1-22.

[135] J. Qi, J. Wu, G. A. Somkuti, J. T. Watson, Determination of the disulfide structure of sillucin, a highly knotted, cysteine-rich peptide, by cyanylation/cleavage mass mapping, *Biochemistry* 40 (2001) 4531-4538.

[136] J. Wu, Y. Yang, J. T. Watson, Trapping of intermediates during the refolding of recombinant human epidermal growth factor (hEGF) by cyanylation, and subsequent structural elucidation by mass spectrometry, *Protein Sci.* 7 (1998) 1017-1028.

[137] J. Qi, W. Wu, C. R. Borges, D. Hang, M. Rupp, E. Torng, J. T. Watson, Automated data interpretation based on the concept of "negative signature mass" for mass-mapping disulfide structures of cystinyl proteins, *J. Am. Soc. Mass Spectrom.* 14 (2003) 1032-1038.

[138] C. R. Borges, J. Qi, W. Wu, E. Torng, A. P. Hinck, J. T. Watson, Algorithm-assisted elucidation of disulfide structure: application of the negative signature mass algorithm to mass-mapping the disulfide structure of the 12-cysteine transforming growth factor beta type II receptor extracellular domain, *Anal Biochem* 329 (2004) 91-103.

[139] J. Wu, J. T. Watson, Optimization of the cleavage reaction for cyanylated cysteinyl proteins for efficient and simplified mass mapping, *Anal. Biochem.* 258 (1998) 268-276.

[140] O. A. Mirgorodskaya, K. F. Haselmann, F. Kjeldsen, R. A. Zubarev, P. Roepstorff, Towards the standard-module approach to disulfide-linked polypeptide nanostructures. I. Methodological prerequisites and mass spectrometric characterization of the test two-loop structure, *Eur. J. Mass Spectrom.* 9 (2003) 139-148.

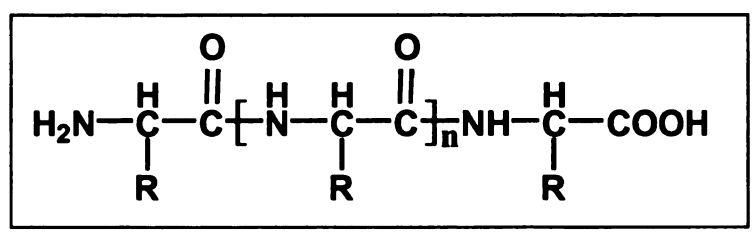
CHAPTER 2. OPTIMIZATION AND STUDY OF EXPERIMENTAL CONDITIONS FOR NUCLEOPHILIC CLEAVAGE OF CYANYLATED CYSTINYL PROTEINS BY PRIMARY AMINES IN DISULFIDE MASS MAPPING METHODOLOGY

1.- INTRODUCTION

1.1 Chemical cleavage of proteins

Proteins and peptides may be cleaved at various amino acid residues by the use of endroproteolytic enzymes (enzymatic cleavage) or by the addition of a chemical reagent (chemical cleavage). Since the early 1960s chemical cleavage methods have been reported for cleaving proteins at specific amino acid residues. Ideally, chemical cleavage methods should have high specificity with minimal degradation of the polypeptide chain. If a particular amino acid residue occurs infrequently and is well positioned along the chain, chemical cleavage at this amino acid residue would generate a relatively small number of peptides of different length that can be purified easily [1].

Proteins are biopolymers composed of 20 different amino acids. The repeating sequence



is held together by amide bonds. Proteins are digested, (hydrolyzed) by specific mammalian enzymes such as trypsin or chymotrypsin; and by less specific plant and animal enzymes, like papain and elastase; or by very aggressive bacterial enzymes (pronase, subtilisin) that hydrolyze all peptide bonds indiscriminately. The chemical cleavage of the peptide group generally requires strong acid or base and elevated temperatures. However, there are some chemical reagents that are able to specifically cleave a protein. For example, cyanogen bromide (CNBr) is able to cleave at the methionine residue with a yield of over 70%; this cleavage reaction is probably the most popular and efficient for proteins [1, 2].

The selective chemical cleavage of the backbone of proteins or peptides can occur at three different sites [3]: the amino-terminal side of the α-carbon atom (p, in Figure 38), the peptide bond (q), or the carboxyl side of the α-carbon (r).

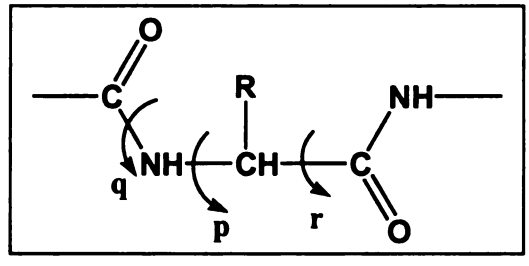
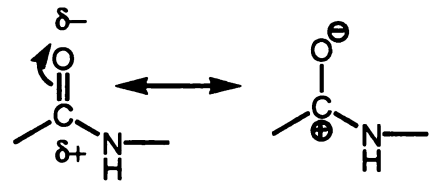


Figure 38. Cleavage sites for a polypeptide chain.

In order to understand the principle of selective chemical cleavage, it is first necessary to understand the general rules of general of hydrolysis of amide bonds (q in Figure 38).

The amide group is a dipole with the negative charge residing on oxygen and the positive charge localized partly on the carbonyl carbon and partly on the amide nitrogen:



To this dipole can be added the hydroxyl anion of a basic medium or the proton of an acid. In both cases, the unstable tetrahedral intermediate breaks down into a carboxylic acid and another amino-terminal residue. There is generally no selectivity in the attack of the H^+ or OH^- on an amide dipole in a protein chain (Figure 39).

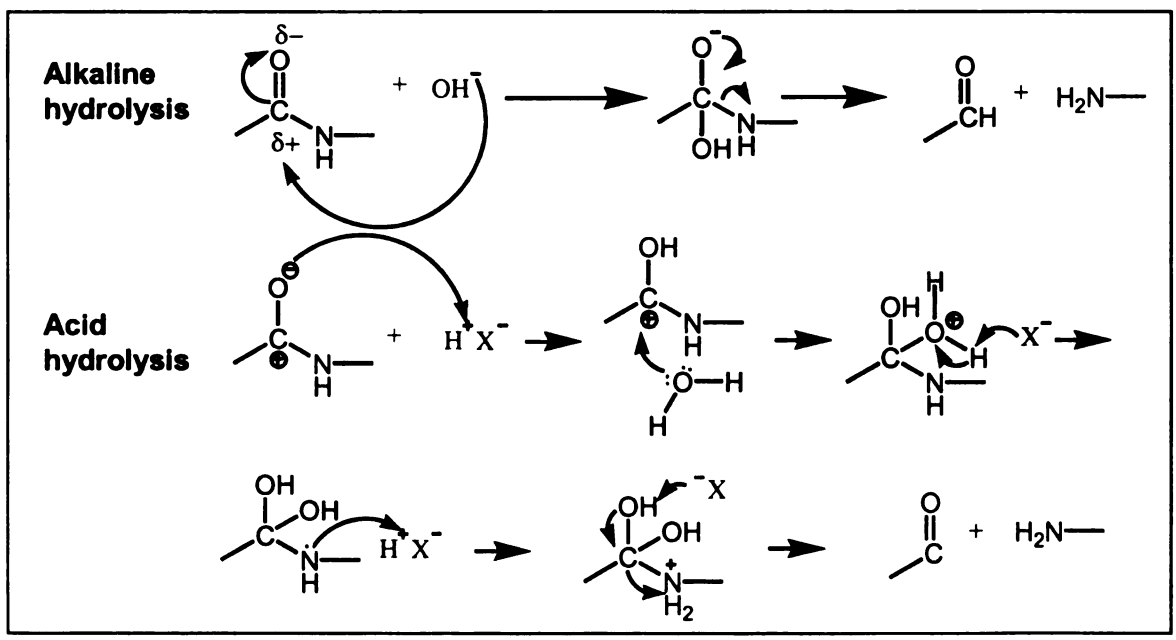


Figure 39. Hydrolysis of amides by acid or base (adapted from [4]).

In selective chemical cleavage, the negative end of the dipole (the imidolate anion) seeks an electrophilic center within the molecule. Such an electrophilic center can be created by a suitable leaving group (X). The departure of the leaving group X need not be in separate discrete steps, but may be concerted. The preferred mechanism is 1,5-interaction, which leads to a thermodynamically favored five-membered ring systems (Figure 40) [4].

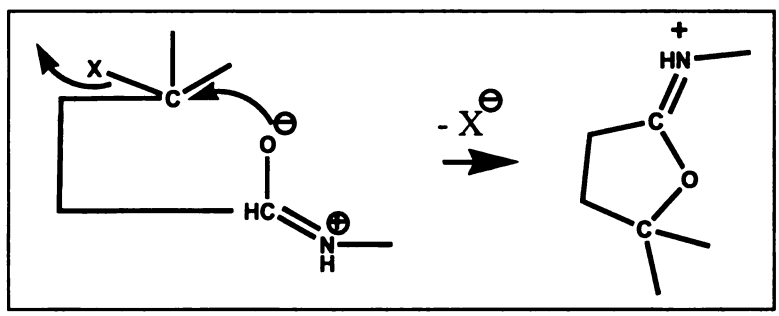


Figure 40. Neighboring group effects associated with chemical cleavage in a protein (adapted from [4]).

The most common site for selective chemical cleavage is the methionine residue [2, 5]. The group in a methionyl peptide that is being removed is the methyl thio-ether. In the first step, the cyano-sulfonium intermediate is formed by interaction with the pseudohalogen cyanogen bromide, and this reaction facilitates removal of sulfur group. The attack of the imidolate anion leads to the unstable imino-γ-lactone, which hydrolyzes to homoserine (lactone) and a new amino-terminal residue. Cleavage of tryptophan, tyrosine, histidine, and reduced phenylalanine follows a similar mechanism (see Figure 41).

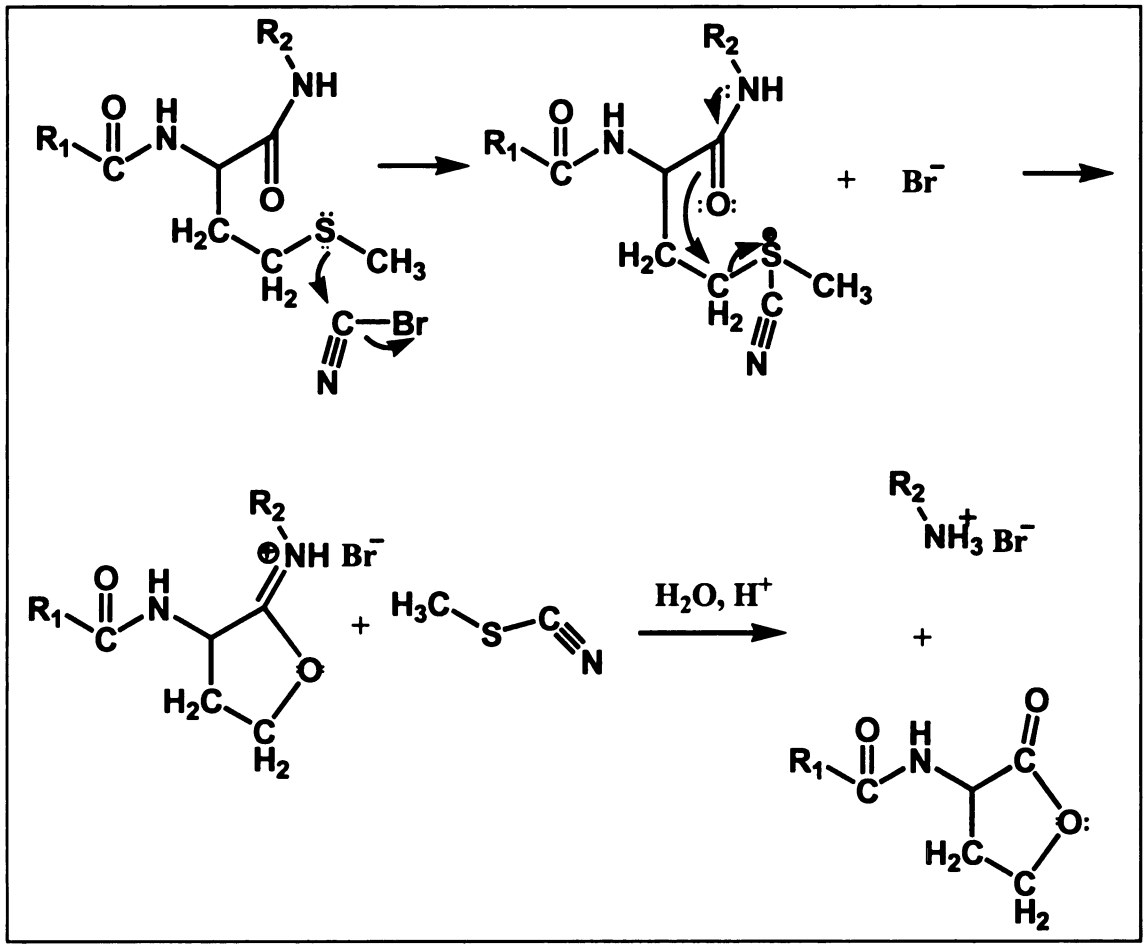


Figure 41. Mechanism of cleavage of methionine peptide bonds with cyanogen bromide (adapted from [1] and [4]).

Other examples of peptide bond cleavage (q in Figure 38) are numerous. In one of these mechanisms, the Nα-amide nitrogen atom can interact with an electrophile position generated at the δ-position on the side chain to form a labile N-acyl–pyrrolidone, –oxazolidone, or –thiazolidine with subsequent cleavage of the acyl-nitrogen function. Another example is the chemical cleavage of the peptide bond on the N-terminal side of cyanocysteine residues (q in Figure 38) under alkaline conditions using ammonia or hydroxide ions as nucleophiles (see Figure 42) [6, 7].

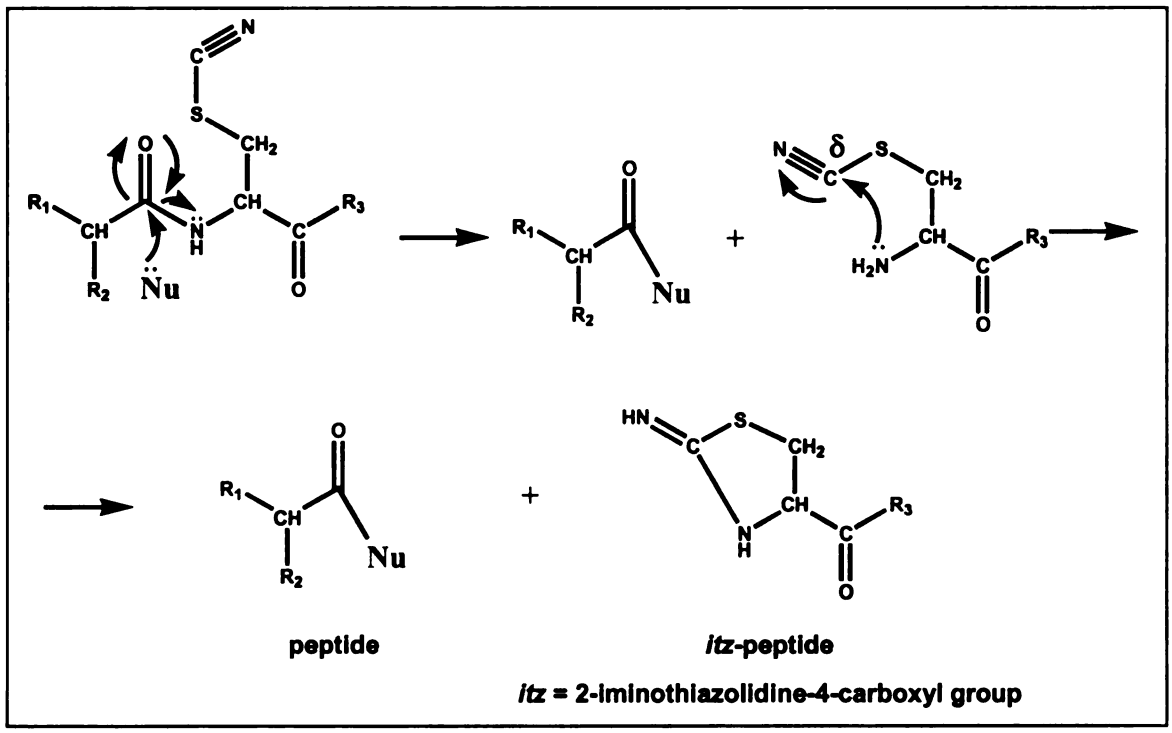


Figure 42. Chemical cleavage of cyanocysteine by means of a nuclephilic attack in alkaline media [8, 9].

1.2 Chemical cleavage in the cyanylation-based disulfide mapping methodology

It was explained in Chapter 1, that during the late '90s, Wu and Watson established a cyanylation (CN)-based chemical method [10] for disulfide mass mapping to overcome many of the shortcomings of the conventional proteolytic approach. This method is based on five main steps: 1) the cystinyl protein is partially reduced, 2) the nascent cysteines produced during reduction of a given disulfide linkage are cyanylated, 3) the polypeptide backbone is then cleaved on the N-terminal side of the modified cysteines, 4) any residual disulfide bonds are reduced to release the cleavage products, and 5) the reduced cleavage reaction

mixture is analyzed by mass spectrometry (MS). Data interpretation leading to the connectivity of cysteines in a disulfide bond relies on recognizing mass spectral peaks that represent CN-induced cleavage products, and mapping these cleavage products to the sequence of the cystinyl protein (see Figure 43).

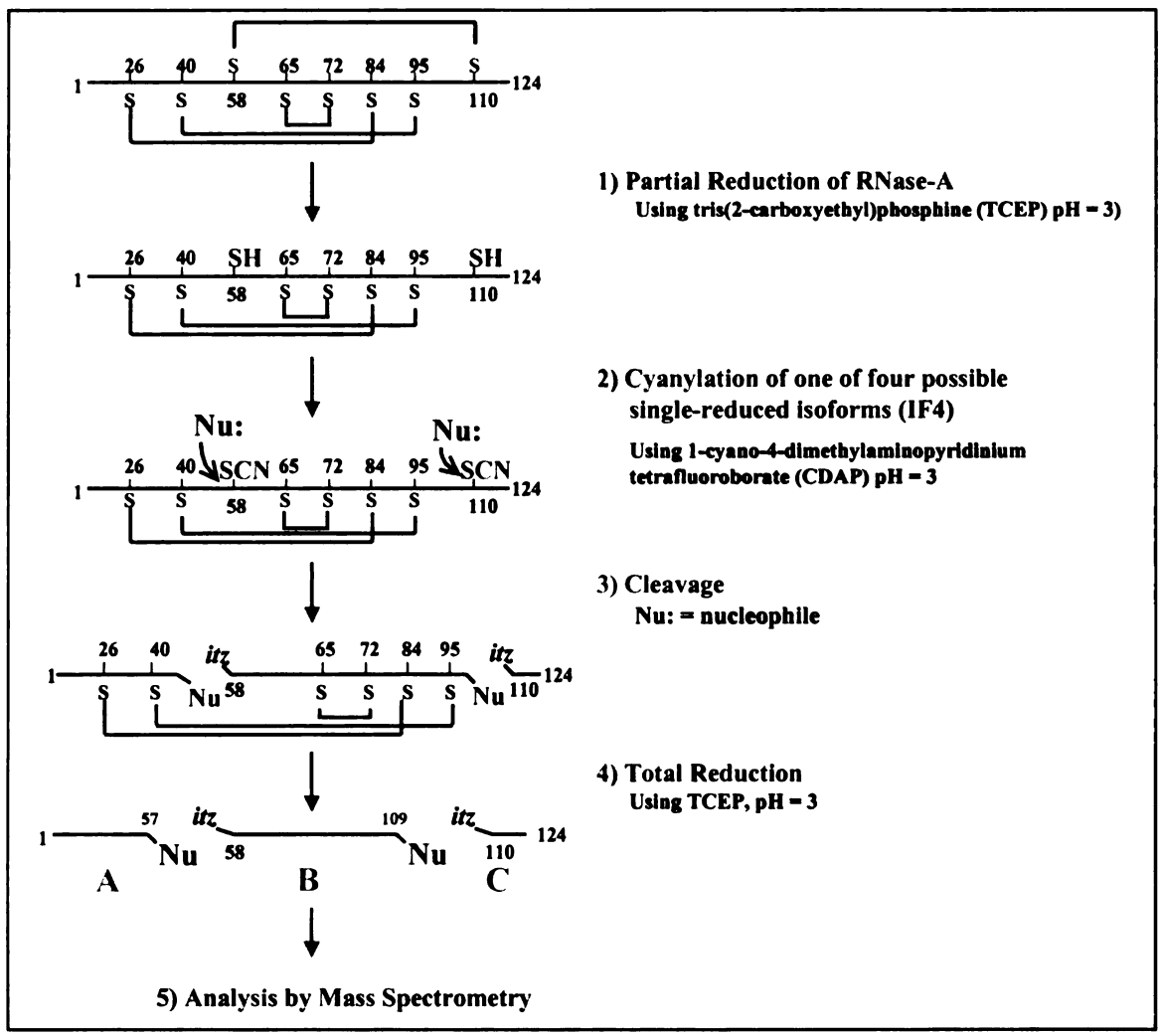


Figure 43. Conceptual representation of the procedure of cyanylation-based disulfide mass mapping.

The cyanylation (CN)-based mass mapping methodology for assignment of disulfide bond connectivities has proven useful in characterizing these linkages even in proteins containing adjacent cysteines [10-12], which is one of the most

challenging analytical problems; studies relating to protein folding have also been carried out using this methodology [13-15].

The essence of the cyanylation-based mass mapping method is based on selective cyanylation of the sulfhydryl group of a cysteine residue, and subsequent nucleophilic attack to promote cyanylation (CN)-induced cleavage of the peptide backbone on the N-terminal side of the modified cysteine as illustrated in Figure 44 [6]. Mass analysis of the cleavage reaction mixture confirms the location of the affected cysteine; e.g., in Figure 44, detection of the fragments consisting of residues 1-9 and *itz*-10-20 (*itz* = iminothiazolidine-blocked N-terminus of a peptide) verifies that a cysteine residue is located at position 10 in the hypothetical 20-residue cystinyl protein. This characteristic cleavage, in particular, has been useful in recognizing the position of free thiols in a study of disulfide isomerization after membrane release in a protein [16].

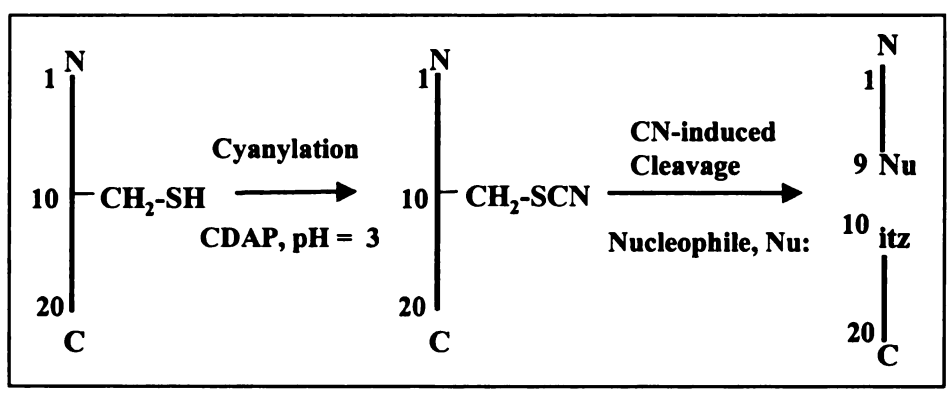


Figure 44. Cleavage of a cyanylated cysteine-containing peptide yields two fragments.

In the CN-based methodology, chemical cleavage of the peptide bond on the N-terminal side of the cyanocysteine is carried out under alkaline conditions, using ammonia or hydroxide ions as nucleophiles [6, 9, 17]. For this cleavage reaction, Jacobson [6] established the first protocol and, Degani and Patchornik [17] studied the pH dependence of the cleavage and β-elimination reactions using di- and tri-peptides as model compounds. Stark [8] summarized the methodology and reviewed some examples and applications. In all this early work, OH^- ion was the nucleophile used under mildly alkaline conditions (pH range: 8-12) with reaction times ranging from 18 to 80 hours. Later, Nakagawa et al [18], applied the CN-cleavage reaction to human parathyroid hormone [hPTH(1-84), containing one cysteine residue at position 35] in order to produce hPTH(1-34). They reported the formation of some by-products during the cleavage step, such as deamidation of asparagine or aspartic acid and methionine oxidation [7, 18, 19].

Wu and Watson [9, 10] systematically studied the rate and yield of the cleavage reaction as a function of pH using OH^- ion or ammonia as nucleophiles. They studied the effects of peptide structure on the production of the desired cleavage products and elimination byproducts, and also found that the presence of proline (and sometimes tyrosine) on the N-terminal side of the cyanylated cysteine plays an important role in promoting the β-elimination side-reaction (see Figure 45). However, a systematic study of the cleavage reaction in a more complex protein has not been reported so far, and some contradictory reports of

using low temperature together with low or high concentrations of the nucleophile have been proposed to maximize the yield of the CN-induced cleavage reaction [7, 19].

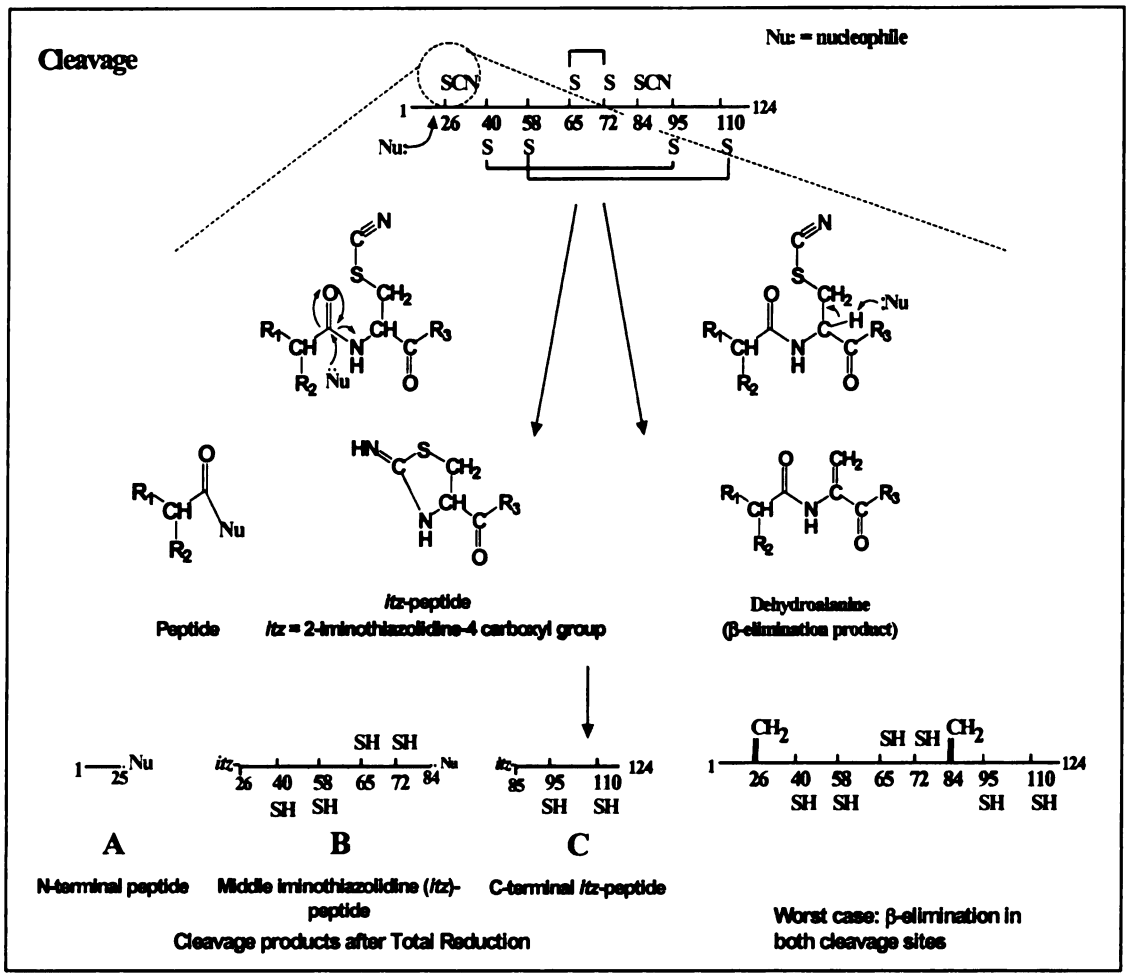


Figure 45. Chemical cleavage in the cyanylation-based disulfide mapping methodology.

In the course of applying the CN-methodology, we have observed occasional disappointing yields of certain cleavage products, the cause of which was apparently related to some complex function of amino acid composition and sequence in the vicinity of the cysteines [9]. Interestingly, in spite of the failure to detect all predicted cleavage products for a given cystinyl protein, we continued

to find good success with the outcome of our analyses. In the interim, we have validated the robust nature of the CN-based methodology through computational analysis of patterns of CN-induced cleavage products [14]. These revealed that certain subsets of the whole array of mapping data are uniquely related to a particular member of the large set of theoretically possible isomeric disulfide structures for a given cystinyl protein. Thus, in principle, detection of only one subset of cleavage products is required to recognize the connectivity of cysteines in a given disulfide [20]. Nevertheless, we would prefer to achieve maximal yield of cleavage products in all cases.

We undertook a systematic study of the nucleophile-mediated CN-induced cleavage reaction using high performance liquid chromatography (HPLC) to monitor the yield of the expected cleavage products as identified by mass spectrometry. The focus of this study was to optimize cleavage of cyanylated cystinyl proteins and to improve the detection of the resulting cleavage products. Primary amines were hypothesized to serve as effective alternative promoters of the cleavage reaction in cyanocysteinyl proteins. This chapter describes results of our study on the feasibility of using a series of primary amines as nucleophiles: methylamine, ethylamine, propylamine, and butylamine. Ammonia was the nucleophile used as a control. The new protocol was applied to analysis of α-lactalbumin from bovine milk type III as well as RNase-A. For a given nucleophile, three main conditions were evaluated: concentration, temperature, and reaction time. Analysis of cleavage reaction mixtures by HPLC produced

chromatograms that made it possible to visualize and analyze changes in the yield of the cleavage reaction under different experimental conditions. Matrix-assisted laser desorption/ionization (MALDI)-MS was used to identify the CN-induced cleavage products.

2.- MATERIAL AND METHODS

2.1 Chemicals

Guanidine hydrochloride was obtained from Boehringer Mannheim Biochemicals, (Indianapolis, IN); 1-cyano-4-dimethylaminopyridinium tetrafluoroborate (CDAP) and tris-(2-carboxyethyl)phosphine (TCEP) hydrochloride were purchased from Sigma Chemical Co. (St. Louis, MO). Methylamine, ethylamine, and propylamine were acquired as hydrochloride salts and butylamine was acquired as a liquid from Aldrich. Ammonia was obtained from EM Science as an aqueous solution containing approximately 14.4 M ammonium hydroxide. Bovine pancreatic ribonuclease A type III-A (RNase-A) and α-lactalbumin from bovine milk Type III were purchased from Sigma Chemical Co. (St. Louis, MO).

Water grade I [21] or better was obtained from a Millipore system, HPLC grade acetonitrile was purchased from EMP chemicals Inc. (Gibbstown, NJ), and spectrophotometric grade trifluoroacetic acid +99% (TFA) was obtained from Aldrich (Milwaukee, WI).

2.2 Mass spectrometry

MALDI mass spectra were acquired on a Voyager DE-STR time-of-flight (TOF) mass spectrometer (Perkin-Elmer Biosystems Inc.) equipped with a 337-nm nitrogen laser. Measurements were made in linear mode with the accelerating voltage set typically to 25,000 V with grid voltage at 95%, guide wire at 0.05%, and extraction delay time at 150 nsec. Time-of-flight to mass conversion was achieved by external and/or internal calibration using standards of bradykinin ($MH^+_{average}$ = 1061.22 Da), bovine pancreatic insulin ($MH^+_{average}$ = 5734.56 Da), horse skeletal myoglobin ($MH^+_{average}$ = 16,952 Da), and/or cytochrome c ($MH^+_{average}$ = 12,361.1 Da) obtained from Sigma (St. Louis, MO). Samples were prepared by spotting 0.5 μL of analyte + 0.5 μL of matrix solution onto a stainless steel sample plate with α-cyano-4-hydroxy cinnamic acid (from Sigma, St. Louis, MO) as the matrix (10 mg/mL) in the modified thin-layer method as described by Cadene and Chait [22]. For some experiments samples were desalted using Zip-Tips from Millipore (Billerica, MA).

2.3 Chromatography

HPLC was performed using a Waters 2695 separation module coupled to a Waters 2487 Dual λ Absorbance detector at 214 nm. The column for purification of isoforms contained a C4 stationaty phase (Vydac # 214TP54, 5-μm particle size, 300 Å pore, 4.6 x 250 mm) and for μ-HPLC separation of reaction mixtures was the column contained a C18 (Vydac # 218TP5115, 5-μm particle size, 300-Å pore, 1.0 x 150 mm). HPLC fractions were collected manually and, if necessary, dried for further use.

2.4 Partial Reduction of RNase-A

The protein sample was placed in a vial where the partial reduction was carried out by adding TCEP in 0.1 M citrate buffer at pH 3.0 for the cysteine content and keeping a concentration of 4 nmol/μL (e.g., 40 nmol of TCEP was reacted with 10 nmol of RNase-A in a total volume of 10 μL) followed by incubation at room temperature for 15 min.

2.5 Cyanylation of single-reduced isoforms of RNase-A

A 20-fold molar excess of 0.1 M CDAP (in 0.1 M Citrate buffer at pH = 3.0) over the total cysteine content was added to the mixture of partially reduced isoforms for the cyanylation step at room temperature for 15 min. Single-reduced

and cyanylated species were purified using HPLC, and then dried. 6M guanidine hydrochloride solution was used to reconstitute the sample in order to obtain (depending on the experiment) solutions between 0.3 and 1 nmol/μL; an accurate concentration was determined, if necessary, by injecting an aliquot of this solution into a μ-HPLC system.

2.6 Cleavage of single-reduced and cyanylated RNase-A isoforms

The nucleophile solution was made using the water grade I or better [21]. Taking pK_a values into account for each reagent, the pH was modified, if necessary, with NaOH, to reach a value of pH = pK_a + 1 to assure predominance of the basic form for the amine (> 90%). Cyanylated isoforms were prepared in 6 M guanidine and treated with a given nucleophile under specified experimental conditions (concentration, temperature, and reaction time). In this report, conventional conditions for the cleavage reaction means that the cyanylated isoform was treated with 1 M ammonia for 60 min at room temperature [10].

2.7 Quenching of the cleavage reaction

After the cleavage reaction was allowed to proceed for the specified time, it was quenched by adding citric acid to establish pH = 3, followed by adding a 0.1 M TCEP solution and incubating at 37°C to completely reduce the remaining disulfide bonds.

2.8 Isolation of single-reduced/cyanylated isoforms of RNase-A by means of high performance liquid chromatography

RNase-A (*m/z* for $MH^+_{average}$ = 13683.3) contains 124 amino acids among which eight cysteines form four disulfide bonds: Cys 26-Cys 84, Cys 40-Cys95, Cys 58-Cys 110, and Cys 65-Cys 72. After partial reduction and cyanylation, single-reduced/cyanylated isoforms of RNase-A can be isolated using reverse phase-HPLC. Because the cyanylated isoforms elute after the intact protein (IP), we refer to them according to their elution order (IF1, IF2, IF3, IF4). The chromatogram in Figure 46 shows the order and nomenclature, which was established during previously reported cyanylation-based mass mapping using MALDI-TOF-MS [10]. Table 2 lists the cysteines that were connected by a disulfide bond in the intact protein, i.e., a disulfide that was reduced to form the particular single-reduced and cyanylated isoform of RNase-A.

Table 2. Nomenclature for the single-reduced isoforms of RNase A.

Single reduced/cyanylated isoform	Reduced Linkage
IF1	Cys 65 – Cys 72
IF3	Cys 26 – Cys 84
IF3	Cys 40 – Cys 95
IF4	Cys 58 – Cys 110

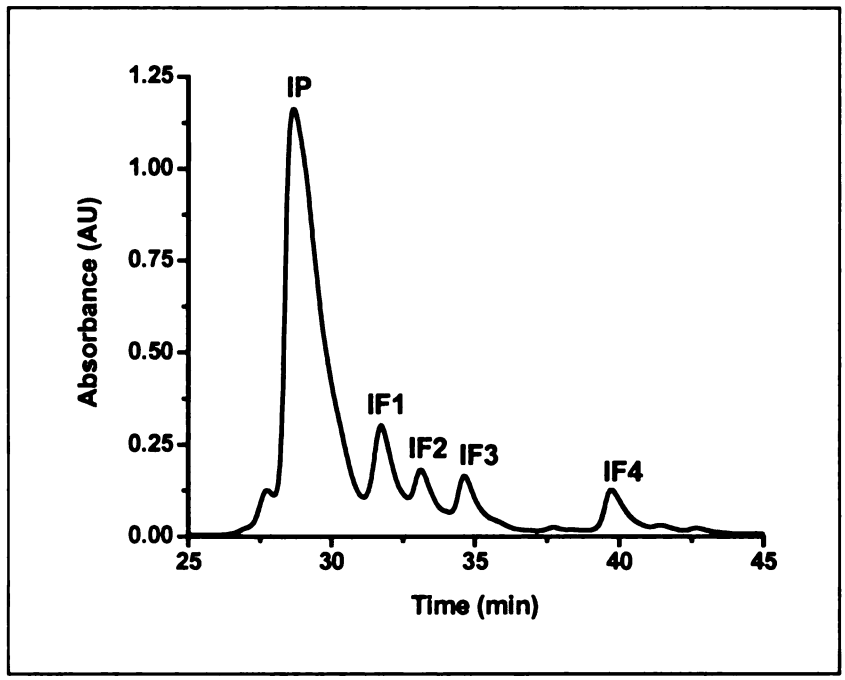


Figure 46. HPLC chromatogram (see text for details) of RNase-A (IP = intact protein) and its single-reduced and cyanylated isoforms (IF1...IF4).

2.9 Partial reduction and cyanylation of α-lactalbumin

α-Lactalbumin (*m/z* for $MH^+_{average}$ = 14,176) contains 123 amino acid residues among which eight cysteines are linked by four disulfide bonds: Cys 6-Cys 120, Cys 28-Cys 111, Cys 61-Cys 77, and Cys 73-Cys 91. Partial reduction and cyanylation of the protein were carried out following a procedure similar to that described for RNase-A. Single-reduced and cyanylated species (IF1L, IF2L, IF3L, IF4L) were separated and collected using μ-HPLC. These species were processed under the optimized cleavage reaction conditions using ammonia, methylamine, or ethylamine as nucleophiles.

3.- RESULTS AND DISCUSSION

3.1 Assessment of nucleophilic cleavage reactions at room temperature

Initial screening experiments showed that yields of cleavage product did not increase significantly over time (from few minutes up to 120 min were evaluated); on the contrary, after 60 min, the yield of some cleavage products decreased, suggesting the existence of secondary reactions. The cleavage reaction for isoform IF4 was observed to be in greater competition with side reactions (data not shown) than those for the other three isoforms. For this reason, we chose IF4 (isolation was explained in section 2.8) as the model isomer for optimizing the cleavage reaction conditions. It was also observed that reaction times less than 30 minutes seemed to be a good choice to avoid decreasing the yield of cleavage products; thus, a series of experiments varying the reaction time was carried out.

At room temperature, IF4 was reacted with 1 M methylamine or 1 M ammonia for different reaction times. Cleavage reaction mixtures were analyzed by μ-HPLC and fractions were collected and analyzed by MALDI-MS for identification. From these chromatograms, it was observed that yields of cleavage products were maximal after 10 minutes of reaction time. In Figure 47, the chromatogram for the cleavage reaction mixture using the conventional

methodology (60 min) is compared with chromatograms of reaction mixtures prepared with methylamine or ammonia for 10 minutes. Ideally, 1 nmol of cyanylated isoform produces 1 nmol of each of the cleavage products, and this result should be reflected in the chromatograms. Using a UV-Vis detector at 214 nm, the signal is proportional to the number of amino acid residues in the cleavage products [23, 24]. Thus, ideally, the peaks for cleavage products A and B should be more intense than that for C, which contains fewer residues (see Figure 43 and values in Table 3); however, it is evident from the chromatograms in Figure 47 that the opposite result was obtained, apparently reaction conditions at room temperature might promote competing side reactions in some cases.

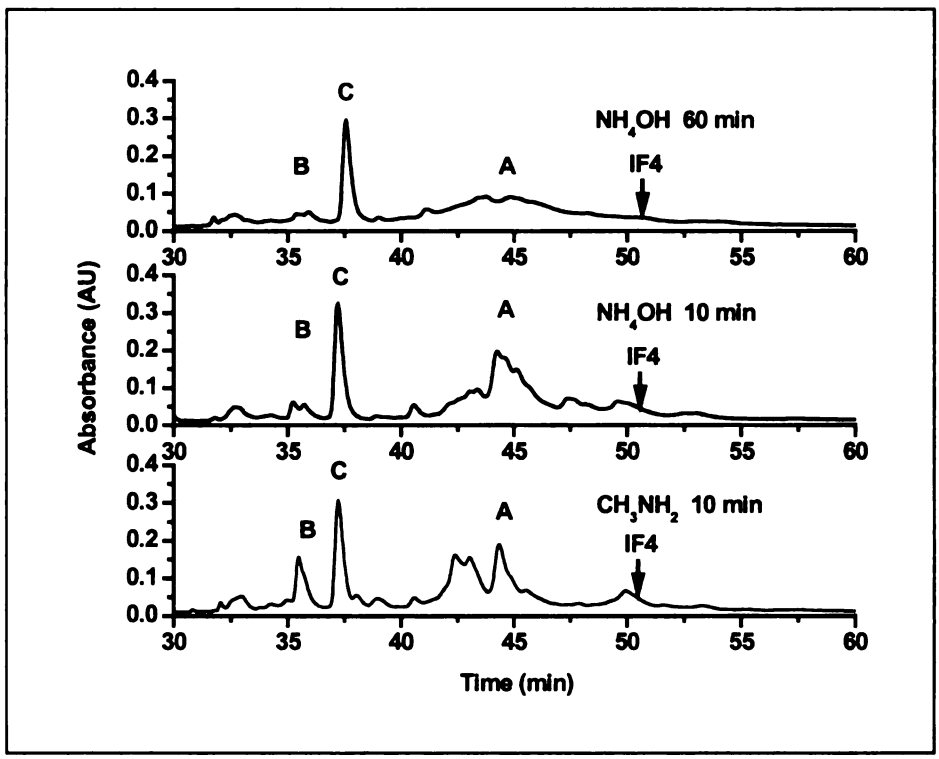


Figure 47. HPLC chromatograms of cleavage products from single-reduced and cyanylated isoform 4 (IF4) of RNase-A treated at room temperature with: Top, 1 M ammonia for 60 min; middle, 1 M ammonia for 10 min; bottom, 1 M methylamine for 10 min. Cleavage products A, B, and C are defined in Figure 43 and in Figure 45. It is also shown the retention time for the intact IF4.

Using methylamine, yields of cleavage products showed a substantial improvement over results using the conventional methodology (e.g. observe the area corresponding to the cleavage fragment B). However, there is still a presence of chromatographic peaks at 42.5 min in Figure 47 corresponding to a mixture of non-specific reaction products near the peak representing cleavage product A. The use of ammonia for only 10 minutes (middle panel of Figure 47) appears to be somewhat better than 60 minutes (top panel of Figure 47) even though products of apparent side reactions still obscure peaks for some of the expected cleavage products. Extended reaction times (beyond 20 minutes) with methylamine resulted in substantial loss of cleavage products A and B, and also some C according to HPLC chromatograms (not shown) compared to that in the lower panel of Figure 47.

Table 3. Calculated *m/z* values for protonated cleavage products of IF4 produced using ammonia or methylamine as nucleophile.

Cleavage product fragment*	Residues	*m/z*	*m/z*
		$-NH_2$	$-NHCH_3$
A	1-57	6351.1	6365.1
B	*itz*-58-109	5766.4	5780.4
C	*itz*-110-124	1659.8	1659.8

MALDI-MS spectra of the cleavage reaction mixture were obtained for samples treated with methylamine for 10 minutes or with ammonia for 60 minutes as shown in Figure 48. As expected, the strength of the mass spectra for certain cleavage products is consistent with the intensity of the corresponding HPLC peaks in Figure 47. This is especially evident for the middle (B) and N-terminal

(A) cleavage products (see Table 3 for *m/z* values) produced using methylamine as the nucleophile. From these results, we concluded that a 10-minute reaction time instead of the previously accepted 60 minutes constituted an analytical advantage. For the other three isoforms (IF1, IF2, and IF3), similar experiments were carried out and, in general, MALDI-MS signals for the cleavage products produced by methylamine were comparable to those produced by using ammonia.

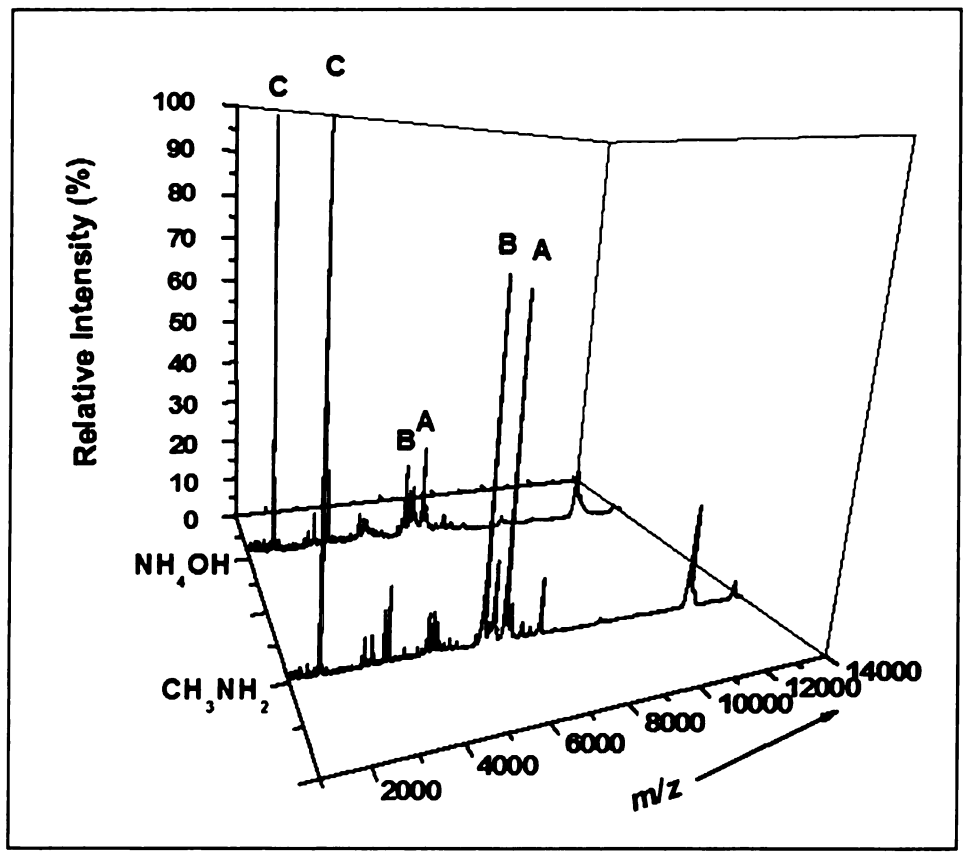


Figure 48. MALDI mass spectra of the cleavage reaction mixture from cyanylated IF4 of RNase-A, using ammonia (60 min) or methylamine (10 min) as the nucleophile. A, B, C are the cleavage products according to Figure 43 and Table 3.

Ethylamine, propylamine, and butylamine were also evaluated for possible improvement in the yield of cleavage products. Because of the hydrophobic

region often found near the cysteine residues, it was thought that the longer hydrophobic alkyl chains of the amine molecules might promote better interaction of the nucleophile with the cyanylated cysteinyl isoforms of RNase-A. However, results from experiments with these longer alkyl amines did not show an advantage over the use of ammonia or methylamine at room temperature for 10 minutes.

3.2 Optimization of experimental conditions for the cleavage reaction using ammonia and methylamine as nucleophiles

3.2a Evaluation of concentration, pH and the use of microwaves at room temperature

Given the experiments described previously, it was necessary to consider if at room temperature there was the possibility for improving the cleavage reaction in order to obtain better yields of cleavage products while minimizing side-reaction products. Figure 49 summarizes results of a series of experiments that were carried out using IF2 as substrate with ammonia or methylamine as the nucleophiles (IF2 was used instead of IF4 to take advantage of available purified isoforms). Thus, because cleavage methodologies have reported using 0.1 M borate buffer at pH 9.0 for up to 18 hours of reaction time [6, 9, 17], a comparison between these experimental conditions and similar conditions using methyl amine or ammonia were carried out. For this comparison, the concentration of

the nucleophile basic form was 1 M at a pH close to the pKa value for the nucleophile in an effort to avoid side reactions at higher pH values (e g. deamidation and beta elimination reactions are more probable). The cleavage reaction was carried out for 12 hours and then chromatograms were obtained from each cleavage reaction mixture. From the chromatograms (not shown), it was evident that the reaction yields were not good using any of the three experimental conditions. Moreover, with the use of ammonia and borate buffer, some intact and only partially degraded IF2 was still observed in the chromatograms, meaning that the cleavage reaction did not proceed completely.

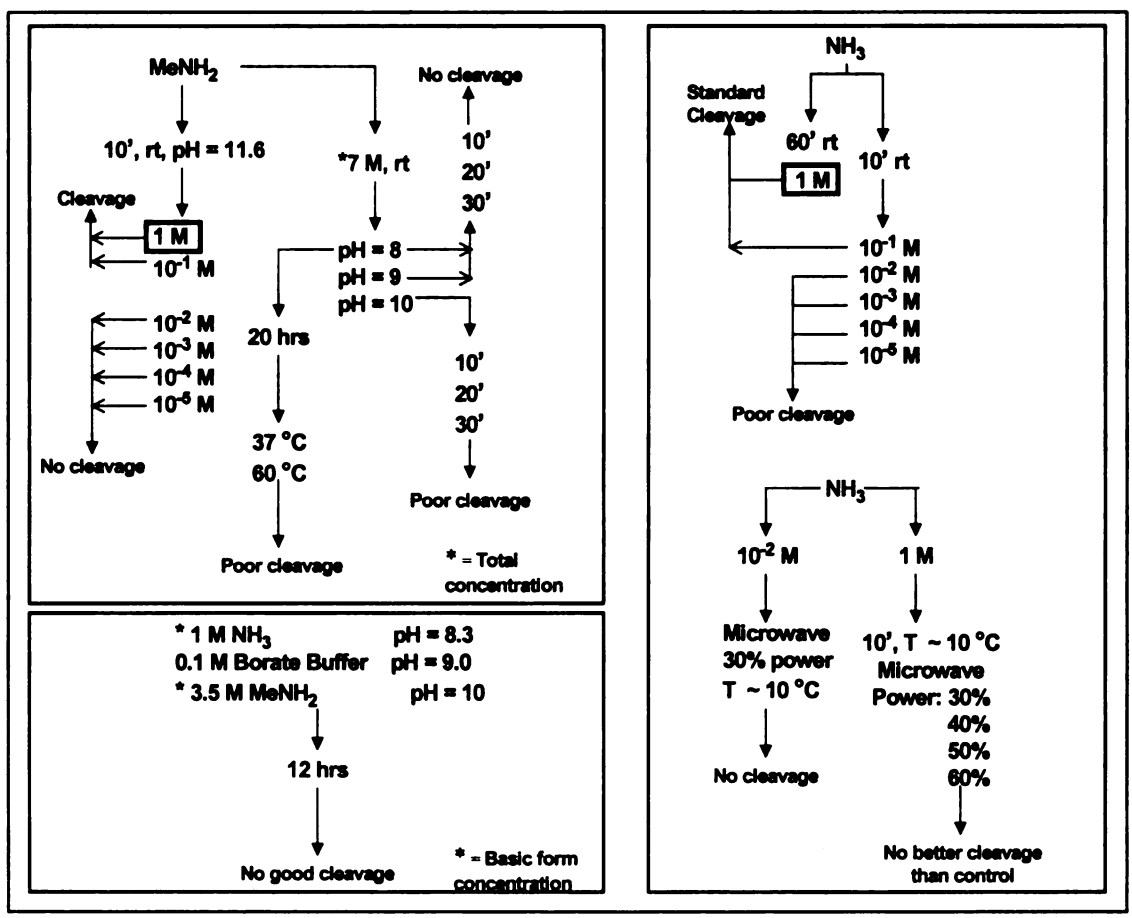


Figure 49. Summary of representative experiments at room temperature using IF2. Squares point out the conditions for the control experiments.

Experiments at lower pH were carried out at different pH values varying temperature and reaction times also in order to check if at values where pH < pKa nucleophiles may promote the cleavage reaction with a good reaction yield. However, only poor cleavage reaction was observed. Similar results were obtained with the use of low nucleophile concentrations. From results of such studies, it was concluded that there was not an advantage to increase the cleavage product yields under those tested experimental conditions.

Microwave-assisted organic synthesis has been applied in a wide variety of reactions [25, 26]; in the protein field it has been applied to shorten acid hydrolysis time for proteins [27, 28]. Also, there are reports of enzymatic digestion assisted by microwaves as an effective and useful method [29-32]. Because of these reports, it was thought that the use of microwaves might be advantageous for the chemical cleavage reaction. Based on results obtained in this laboratory for other studies using microwaves in protein analysis, a series of experiments was carried out using a laboratory-adapted microwave oven. However, there was no significant improvement in the yield of cleavage products; thus, further use of microwaves was discontinued.

3.2b Effect of low temperature and reaction time on the cleavage reaction

Performance of the cleavage reaction at 2 °C results in a significant diminution in the products of side-reactions as shown in the comparison of chromatograms in Figure 50. This observation is consistent with results from earlier studies of the cleavage reaction of some peptides or proteins [18, 33]. In our study of the effect of reduced temperature on the CN-induced cleavage reaction of a more complex protein, IF4 was treated with 1 M methylamine or 1 M ammonia at 2 °C in a cold room. A noticeable improvement in the yield of the CN-induced cleavage product was realized at 2 °C when methylamine was the nucleophile for a 10-min reaction (see Figure 50). In comparison, the area under the curve for the middle cleavage product B at 2 °C is about 25% greater than that for the equivalent reaction carried out at room temperature. Similarly, the cleaner and significantly different HPLC peak for the cleavage product A in the lower panel of Figure 50 suggests reduced competition from side reactions at lower temperature. We also investigated the influence of longer reaction time for samples treated with methylamine at low temperature.

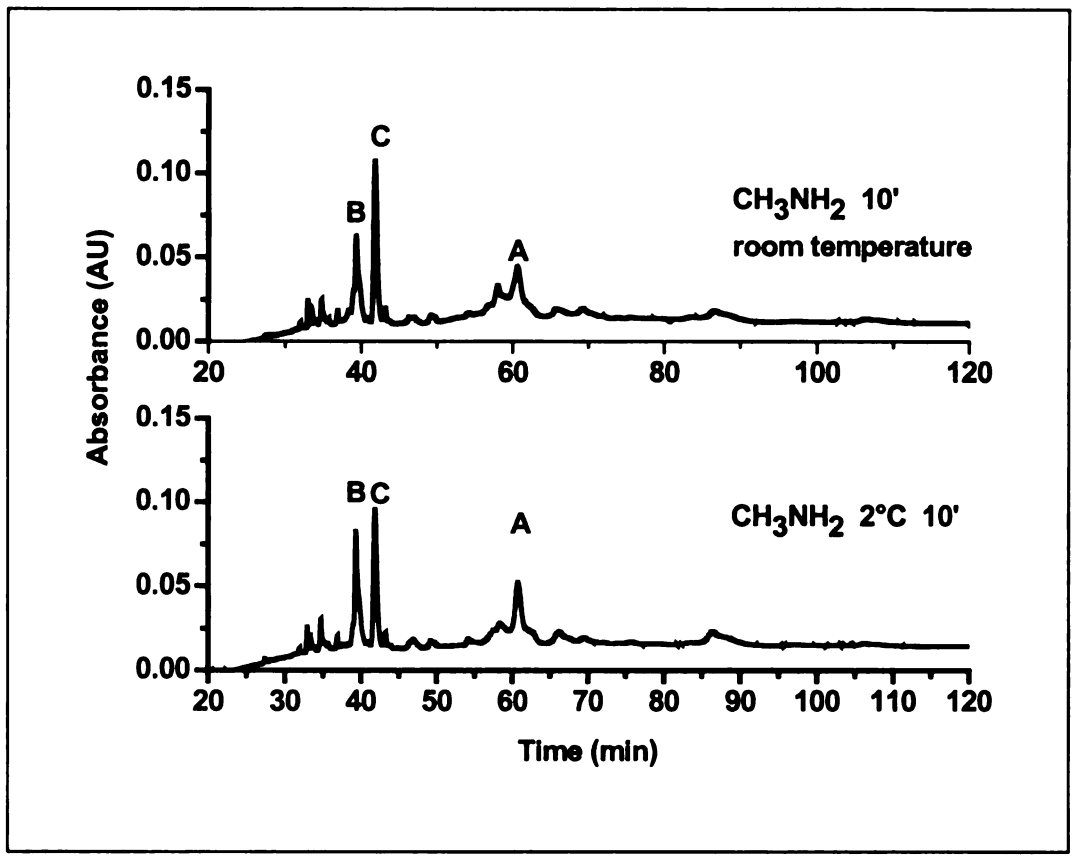


Figure 50. Comparison of HPLC chromatograms of cleavage reaction mixtures of IF4 using 1 M methylamine at room temperature or at 2 °C for 10 minutes. The peaks labeled A, B, and C represent the cleavage products as defined in Figure 43.

The results (data not shown) indicated no appreciable increase in the yields of the CN-induced cleavage products. Hence, 10 minutes at 2°C seems to be optimal in maximizing the yield of cleavage products, while minimizing side-reactions.

Samples treated with 1 M ammonia at low temperature for 10 minutes showed only a moderate improvement in the yield of cleavage products over those obtained using previously described conditions [9, 10]. Additional experiments with longer reaction times in 1 M NH_3 showed only a slight

improvement in yield of the CN-induced cleavage fragments, but less than the yields using 1 M methylamine (data not shown).

3.2c Effect of nucleophile concentration on the cleavage reaction

The next factor evaluated was the concentration of the nucleophiles (at pH ~ 11.6). Samples of IF4 were treated with different concentrations of ammonia or methylamine, and the cleavage reaction mixtures were analyzed using μ-HPLC. The cleavage product yield increased linearly with [NH_3] up to about 5 M and with [CH_3NH_2] up to 2 M. These concentrations were considered optimum because, in the case of ammonia beyond 5 M, there is not a noticeable increment (<10%) of the cleavage product yields that could justify an increase in the nucleophile concentration; for methylamine, at concentrations higher than 2 M, the yield of some cleavage products decreased, suggesting the involvement of side reactions (Figure 51).

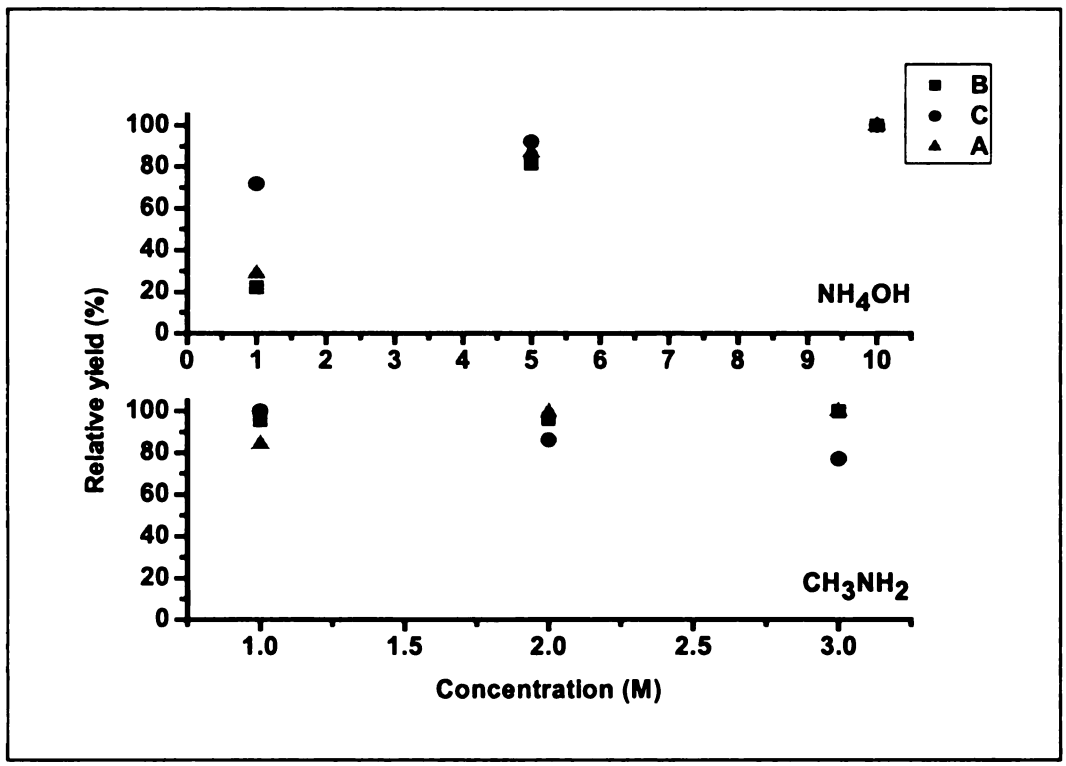


Figure 51. Relative yield of cleavage products using methylamine for IF4 from RNase-A as function of the concentration at 2°C. Labels B, C, and A correspond to the cleavage fragments according Table 3 and Figure 43.

3.2d Repeatability of cleavage reaction yields

Because it is important to achieve statistical control of a measurement [34], a series of analyses was conducted to evaluate the precision of the cleavage reaction yield. Samples at three isoform concentration levels were selected for the experiments: ~0.022 nmol/μL (denoted 1x), ~0.103 nmol/μL (5x), and ~0.340 nmol/μL (15x); three reaction conditions for each level were studied: 5 M ammonia (5000 nmol/μL) at 2 °C for 10 minutes, 2 M methylamine (2000 nmol/μL) at 2 °C for 10 minutes, and the conventional conditions previously

described in section 2.6 and elsewhere [9, 10]. Triplicate analyses were carried out for each level.

The cleavage reaction mixtures were analyzed by HPLC with UV absorption for determination of the yield of each cleavage product. Because UV absorptivity is proportional to the number of peptide bonds in a protein or peptide [23, 35], calibration curves were made using the peak area from chromatograms for known concentrations of peptides containing a specified number of amino acid residues. From the peak areas in chromatograms for the separation of the cleavage reaction mixture, cleavage product amounts and yields were calculated for each experiment taking into account that ideally 1 nmol of IF4 would produce 1 nmol of each cleavage fragments.

Statistical data were obtained for isoform levels 1x, 5x, and 15x using the JMP 5.1© statistical software [36]. The use of this software allows graphical visualization to compare multiple data groups (Figure 52).

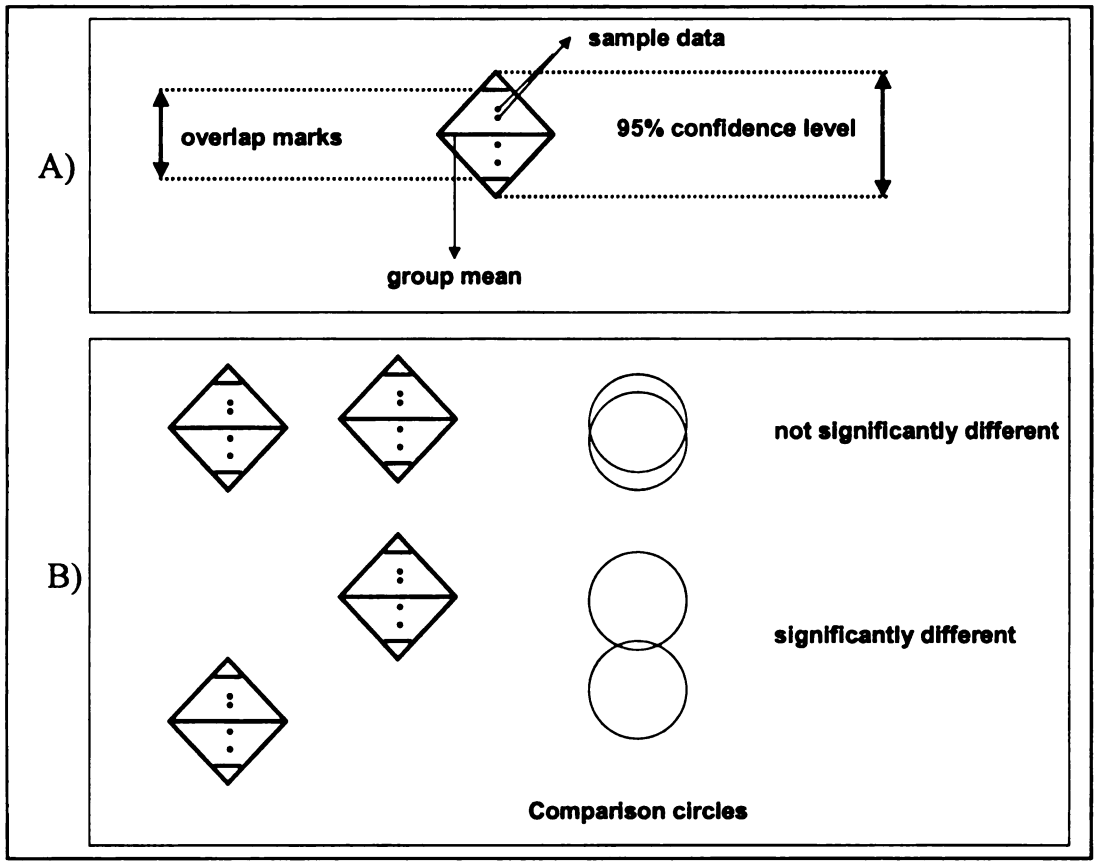


Figure 52. A) Mean diamonds graphs. The middle line in the diamond represents the mean for a group of data. The vertical endpoints form the 95% confidence interval for the mean. Overlap marks are useful for comparison of means, but it is better to use of comparison circle plots. B) Two hypothetical cases where means of two sample data groups are *not significantly* and are *significantly* different. Comparison circle plots for each group are evident each case.

These statistical results indicate that similar treatments produce similar amounts of cleavage products proportional to the initial amount of the isoform; e.g., the amount of cleavage products at level 15x was roughly 15 times the amount at level 1x under identical experimental conditions. In general, there was no significant difference ($p = 0.05$) in the proportional amount of cleavage products at different levels. It is evident that the molar ratio of nucleophile/isoform is so high that a change in the amount of available isoform for the reaction does not affect the yield of cleavage fragments significantly. The

coefficients of variance (CV) at levels 15x and 5x are, in general, no higher than 10% (see Table 4 and Figure 53). For level 1x, the CV values are relatively high (>10%) mainly because the chromatographic peaks are smaller and the chromatographic integration is more affected by the background noise.

Comparison of mean values for the amount of the cleavage products obtained shows that the cleavage reaction yields for fragments A and B vary significantly with reaction conditions (see Table 4); the yield of these cleavage fragments is much higher under the optimized methodology. It is not surprising that the yield of cleavage fragment C did not change significantly with reaction conditions because according to our previous studies, this fragment is produced with great facility. For example, the yield of fragment C is the same at 10 minutes under the optimized cleavage conditions as it was after using the conventional conditions.

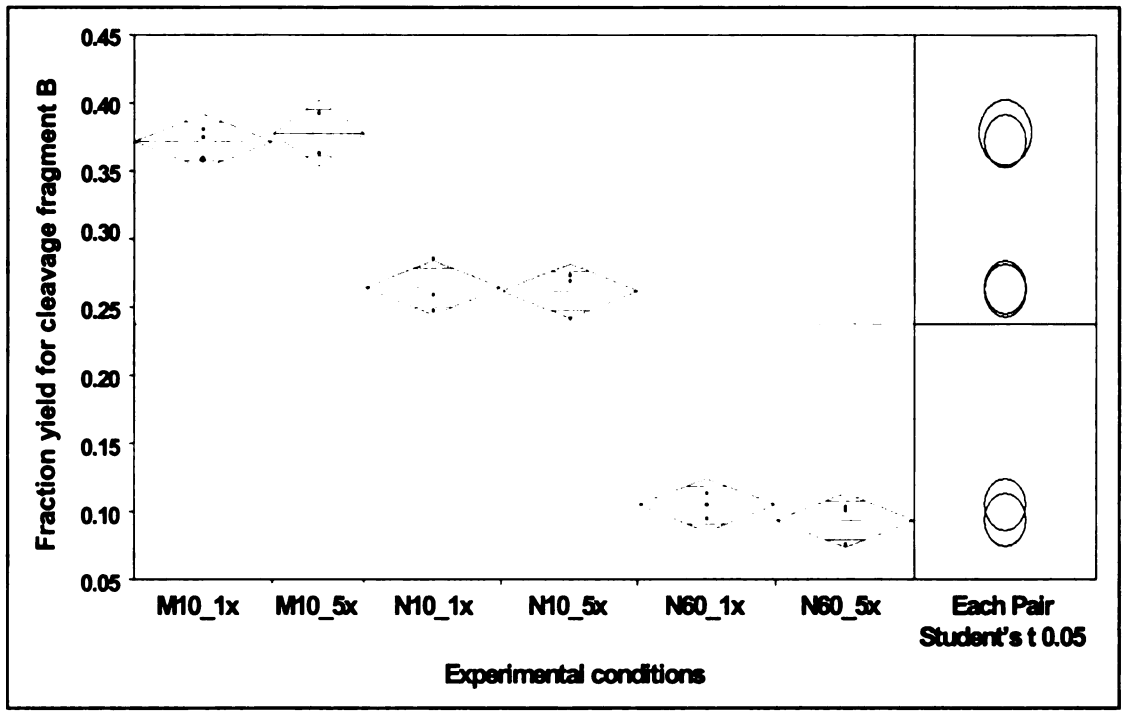


Figure 53. Distribution analysis and mean comparison for yield of cleavage product B. M10 = Experiments with methylamine, 10 minutes at 2 °C. N10 = Experiments with ammonia, 10 minutes at 2 °C. N60 = Experiments with ammonia, 60 minutes at room temperature.

Table 4. Representative data for cleavage product yields for IF4 obtained at level 15x under different experimental conditions. In parenthesis, the coefficient of variation (CV) is expressed as a percentage, for experiments under the same treatment.

15x	A yield in percent (CV)	B yield in percent (CV)	C yield in percent (CV)
$MeNH_2$, 2 °C, 10'	41 (10.8%)	30 (4.8%)	89 (4.9%)
NH_3, 2 °C, 10'	27 (4.4%)	21 (3.7%)	101 (2.14%)
NH_3, rt, 60'	2 (4.9%)	6 (9.6%)	95 (5.0%)

rt = room temperature

3.3 Evaluation of the optimized experimental conditions

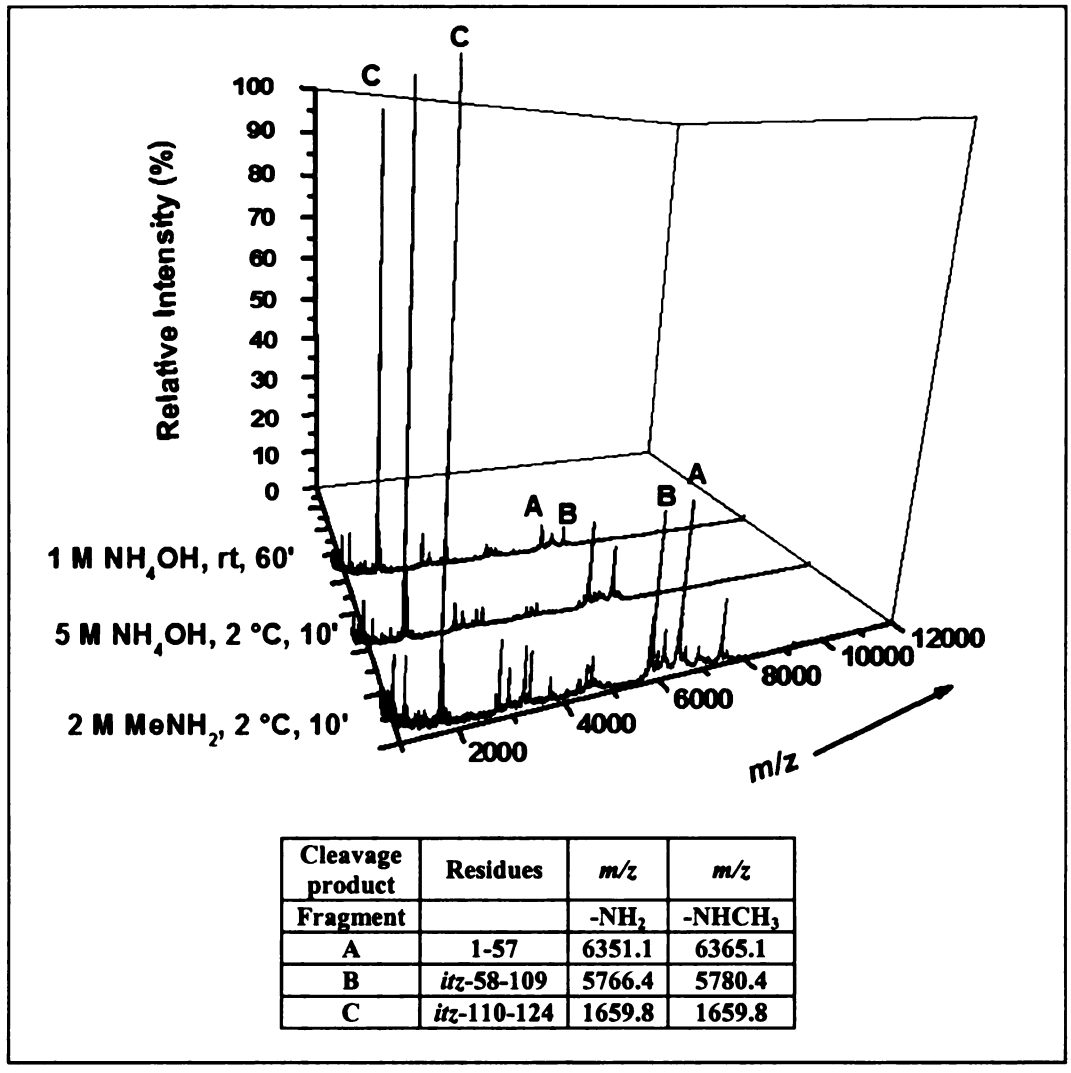
3.3a Comparison of MALDI-mass spectra of cleavage reaction mixtures prepared under modified conditions at low temperature

Our main goal is to determine the disulfide linkage in a cystinyl protein through the correct assignment and identification of the CN-induced cleavage products by using mass spectrometry. Better yields of cleavage products lead to better quality mass spectra, and by consequence, more precise and unambiguous assignment of cleavage fragments. In order to evaluate mass spectra of cleavage reaction mixtures prepared under the newly optimized experimental conditions, single-reduced and cyanylated isoforms of RNase-A (IF1 to IF4) were treated with methylamine, ammonia, or other candidate nucleophiles including: ethylamine, propylamine, and butylamine. For a controlled comparison, a sample of IF4 was also treated under the conventional methodology (1 M ammonia, 60' reaction time). Samples of cleavage reaction mixtures were spotted on a plate for the MALDI-MS experiment according to the method described in the experimental section.

Figure 54 shows a comparison of MALDI-mass spectra for the cleavage reaction mixture of IF4 using CH_3NH_2 or NH_3 under different experimental conditions. It is evident that peaks labeled A and B, representing the N-terminal and middle fragments, respectively, become more prominent with 5 M NH_4OH

and 2 M CH_3NH_2 compared to those observed in the spectrum of the reaction mixture prepared under the traditional experimental conditions (60 min at room temperature in 1 M of NH_4OH).

Expected and characteristic is the shift in the mass of the cleavage products A and B (as defined in Figure 43) according to the nucleophile used. These two cleavage products terminate in the form of an amide that incorporates the nucleophile (Nu). For example, if the nucleophile is ammonia, Nu is $-NH_2$; in the case of methylamine, Nu is $-NH-CH_3$, etc. Thus, cleavage products A and B generated with a mixture of ammonia and methylamine as nucleophiles will be represented by a pair of mass spectral peaks separated by 14 *m/z* units. On the other hand, the *itz*-C-terminal product (C) will always show the same molecular mass regardless of the nucleophile used.



Cleavage product	Residues	*m/z*	*m/z*
Fragment		$-NH_2$	$-NHCH_3$
A	1-57	6351.1	6365.1
B	*itz*-58-109	5766.4	5780.4
C	*itz*-110-124	1659.8	1659.8

Figure 54. MALDI mass spectra of the reaction mixture resulting from the cleavage of cyanylated IF4 from RNase-A by the indicated nucleophiles. Peaks labeled A, B, and C represent segments of the sequence as defined in Figure 43.

Analytical advantage can be gained in the use of a binary mixture of homologous nucleophiles for cyanylation-induced cleavage of partially reduced isoforms of cystinyl proteins. The appearance of pairs of mass spectral peaks offset by 14 (or 28, 42, etc) mass units facilitates recognition of diagnostic CN-induced cleavage products, and increases confidence in results when samples present weak signals or interfering chemical noise during analysis by MS. The

simultaneous use of two homologous nucleophiles is also applicable to the identification of cleavage products from double-reduced isoforms, which are also useful in the elucidation of disulfide linkages [20].

Figure 55 shows a comparison of the MALDI mass spectra for the cleavage reaction mixture of IF4 treated with different primary amines. Mass spectral peaks representing cleavage products from reactions with methylamine show a higher intensity compared to those from analysis of reaction mixtures processed with the other amines.

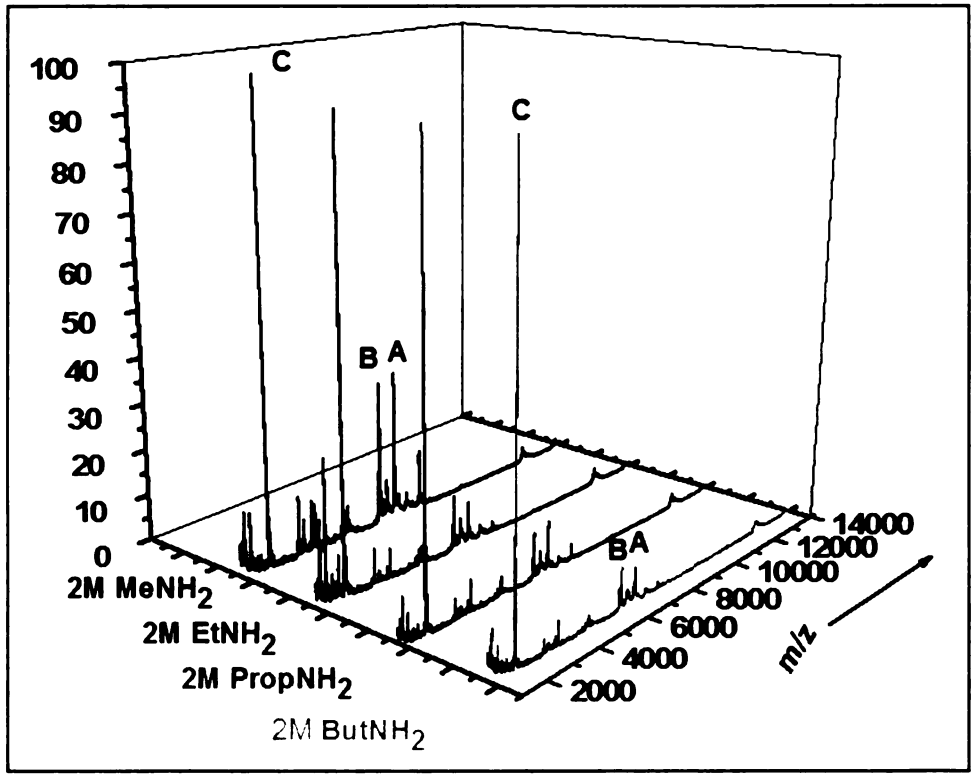


Figure 55. MALDI-MS spectra of the cleavage product mixture resulting from the processing of IF4 from RNase-A with primary amines at low temperature (2 °C) and a concentration of 2 M. From back to the front: methylamine, ethylamine, propylamine, or butylamine.

It has been demonstrated [9] that amino acids with bulky side chain groups, such as Pro and Tyr, on the N-terminal side of the cyanylated cysteine inhibit the cleavage process. The alkyl chain of the alkyl amine increases the nucleophilicity, but steric problems can be present with larger hydrocarbon chains. According to our observations, methylamine produces a better cleavage product yield because it is a good nucleophile, but it has a relatively small methyl group. In the case of the other primary amines, even though they could also be considered good nucleophiles [37, 38], the length of the hydrocarbon chain seems to impair their effectiveness during the nucleophilic attack. However, there is no evidence for any specific side reaction. Therefore, methylamine appears to be a good compromise between functioning as a strong nucleophile and avoiding steric hindrance.

3.3b Application of the modified methodology to the analysis of α-lactalbumin

In order to evaluate the modified methodology with a different protein, α-lactalbumin (*m/z* for $MH^+_{average}$ = 14176) was used as another model protein; it contains 123 amino acid residues among which eight cysteines are linked by four disulfide bonds: Cys 6-Cys 120, Cys 28-Cys 111, Cys 61-Cys 77, and Cys 73-Cys 91. Partial reduction and cyanylation of α-lactalbumin were carried out following a procedure similar to that described for RNase-A. Single-reduced and cyanylated isoforms of α-lactalbumin were separated and collected using reverse

phase-HPLC. Each isoform was dried, reconstituted in 6 M guanidine, and used to prepare samples for treatment with different nucleophiles: ammonia, methylamine, or ethylamine under the optimized experimental conditions. MALDI mass spectra of reaction mixtures under optimized conditions were compared to those previously reported using the conventional methodology [9, 10].

Although MALDI-MS is not a quantitative technique, relative peak intensities for the cleavage products might be used as a rough estimate of the cleavage product yields. Mass spectral peaks were more discernible during analysis of cleavage reaction mixtures prepared with ammonia under optimized reaction conditions (5 M NH_4OH, 2 °C, 10 minutes) than in those prepared with the conventional methodology; thus, we can infer that yields of the cleavage products were better under the optimized reaction conditions with ammonia.

The pattern of cleavage products changed when using methyl- or ethylamine as the nucleophile in that cleavage fragment 1-60 was the most abundant, whereas cleavage fragment *itz*-77-123 was the most abundant when using ammonia (see Table 5). These observations suggest that it is easier for the primary amines to cleave the cyanylated cysteine located at position 77 than the one located at position 61. This phenomenon can be rationalized by examining the amino acid sequence for α-lactalbumin in which the amino acid residue on the N-terminal side of cysteine 61 is Trp, the bulky side chain of which may present steric hindrance to the cyclization cleavage reaction on the N-

terminal side of the cyanylated cysteine residue. This phenomenon may be analogous to that reported previously [9] for Pro or Tyr. It also seems reasonable that the sidechain of Trp might interfere with access of the nucleophile to the cyanylated cysteine.

It was explained earlier that cleavage product B of a given single-reduced isoform, obtained by reaction with methylamine, is 14 Da heavier than that obtained with ammonia. This feature made it possible to recognize the cleavage product *itz*-73-90 from one of the isoforms of α-lactalbumin even though the relative intensity in the MALDI mass spectrum (data not shown) for this fragment was low (<1%).

Table 5. Relative intensities from MALDI mass spectra for protonated cleavage products produced using ammonia, methylamine, or ethylamine as cleavage reagents for single-reduced/cyanylated α-lactalbumin isoforms (IF1L, IF2L, IF3L, IF4L). Calculated *m/z* values are for protonated cleavage products produced by nucleophilic attack by ammonia.

IF1L Open Cys 6 – Cys 120		**Nucleophile**			
Fragment	***m/z***	***NH_3OH, 1 M, rt, 60'**	**NH_3OH 5 M, 2 °C, 10'**	**CH_3NH_2 2 M, 2 °C, 10'**	**$CH_3CH_2NH_2$ 5 M, 2 °C, 10'**
1 – 5	618.7	ND	ND	ND	ND
itz-6 – 119	13, 135.7	ND	2.91	1.61	2.18
itz-120 – 123	517.6	100	100	100	100
IF2L Open Cys 61 – Cys 77		**Nucleophile**			
Fragment	***m/z***	***NH_3OH, 1 M, rt, 60'**	**NH_3OH 5 M, 2 °C, 10'**	**CH_3NH_2 2 M, 2 °C, 10'**	**$CH_3CH_2NH_2$ 5 M, 2 °C, 10'**
1 – 60	6918.7	~ 8	100	76.0	65.1
itz-61 – 76	1800.8	~ 8	26.2	45.8	21.3
itz-77 – 123	5552.5	100	69.7	100	100
IF3L Open Cys 28 – Cys 111		**Nucleophile**			
Fragment	***m/z***	***NH_3OH, 1 M, rt, 60'**	**NH_3OH 5 M, 2 °C, 10'**	**CH_3NH_2 2 M, 2 °C, 10'**	**$CH_3CH_2NH_2$ 5 M, 2 °C, 10'**
1 – 27	3125.6	100	100	100	51.0
itz-28 – 110	9525.6	~ 2	2.01	1.99	2.10
itz-111 – 123	1620.8	~ 98	98.5	84.4	100
IF4L Open Cys 73 – Cys 91		**Nucleophile**			
Fragment	***m/z***	***NH_3OH, 1 M, rt, 60'**	**NH_3OH 5 M, 2 °C, 10'**	**CH_3NH_2 2 M, 2 °C, 10'**	**$CH_3CH_2NH_2$ 5 M, 2 °C, 10'**
1 – 72	8258.1	~ 5.00	7.86	12.9	5.18
itz-73 – 90	2098.3	~ 4.00	7.08	4.10	1.00
itz-91 – 123	3915.6	100	100	100	100

ND = Not detected, rt = room temperature, * Data from reference [10]

4.- CONCLUSIONS

Methylamine (2 M) was found to be superior to ammonia as the cleavage nucleophile in promoting the CN-induced cleavage reaction during disulfide mass mapping. However, increasing the concentration of ammonia to 5 M (beyond the previously used 1 M) in combination with low temperature (2° C instead of room temperature) and short reaction time increased the cleavage product yields substantially (in some cases, by an order of magnitude). Low temperature (2 °C) and a short reaction time (10 min) minimize alkali-induced damage of the cleavage products and other side reactions that occur during longer incubation times in the original protocol (1 hour or more). Good repeatability (less than 10% relative standard deviation) was observed in the yield of cleavage products from ~100-1000 pmol of starting material.

The use of two or more homologous nucleophiles can be visualized as an additional tool to facilitate the identification of cleavage products from mass spectra. Representation of certain cleavage products by pairs of mass spectral peaks separated by 14 mass units helps the analyst to distinguish these structurally diagnostic species, especially in cases of low intensity signals and/or in the presence of potentially interfering chemical noise.

5.- REFERENCES

[1] K. K. Han, C. Richard, G. Biserte, Current developments in chemical cleavage of proteins, *Int. J. Biochem.* 15 (1983) 875-884.

[2] B. J. Smith, Chemical cleavage of polypeptides, *Methods Mol. Biol.* 211 (2003) 63-82.

[3] T. F. Spande, B. Witkop, Y. Degani, A. Patchornik, Selective cleavage and modification of peptides and proteins, *Adv. Protein Chem.* 24 (1970) 97-260.

[4] B. Witkop, Chemical cleavage of proteins. Selective fragmentations and modifications reveal structure, *Science* 162 (1968) 318-326.

[5] S. Doonan, Peptides and proteins, Royal Society of Chemistry, Cambridge, UK., 2002.

[6] G. R. Jacobson, M. H. Schaffer, G. R. Stark, T. C. Vanaman, Specific chemical cleavage in high yield at the amino peptide bonds of cysteine and cystine residues, *J. Biol. Chem.* 248 (1973) 6583-6591.

[7] S. Nakagawa, Y. Tamakashi, T. Hamana, M. Kawase, S. Taketomi, Y. Ishibashi, O. Nishimura, T. Fukuda, Chemical cleavage of recombinant fusion proteins to yield peptide amides, *J. Am. Chem. Soc.* 116 (1994) 5513-5514.

[8] G. R. Stark, Cleavage at cysteine after cyanylation, *Methods Enzymol.* 47 (1977) 129-132.

[9] J. Wu, J. T. Watson, Optimization of the cleavage reaction for cyanylated cysteinyl proteins for efficient and simplified mass mapping, *Anal. Biochem.* 258 (1998) 268-276.

[10] J. Wu, J. T. Watson, A novel methodology for assignment of disulfide bond pairings in proteins, *Protein Sci.* 6 (1997) 391-398.

[11] J. Wu, J. T. Watson, Assignment of disulfide bonds in proteins by chemical cleavage and peptide mapping by mass spectrometry, *Methods in Molecular Biology (Totowa, NJ, United States)* 194 (2002) 1-22.

[12] J. Qi, J. Wu, G. A. Somkuti, J. T. Watson, Determination of the disulfide structure of sillucin, a highly knotted, cysteine-rich peptide, by cyanylation/cleavage mass mapping, *Biochemistry* 40 (2001) 4531-4538.

[13] J. Wu, Y. Yang, J. T. Watson, Trapping of intermediates during the refolding of recombinant human epidermal growth factor (hEGF) by cyanylation, and subsequent structural elucidation by mass spectrometry, *Protein Science* 7 (1998) 1017-1028.

[14] J. Qi, W. Wu, C. R. Borges, D. Hang, M. Rupp, E. Torng, J. T. Watson, Automated data interpretation based on the concept of "negative signature mass" for mass-mapping disulfide structures of cystinyl proteins, *J. Am. Soc. Mass Spectrom.* 14 (2003) 1032-1038.

[15] W. Wu, J. T. Watson, F. J. Stevens, R. Yousefzai, L. Anderson, Chloroplast fructose bisphosphatase contains alternate disulfide bonds, *Abstracts, 35th Great Lakes Regional Meeting of the American Chemical Society, Chicago, IL, United States, May 31-June 2* (2003) 110.

[16] M. Xu, A. Arulandu, D. K. Struck, S. Swanson, J. C. Sacchettini, R. Young, Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme, *Science* 307 (2005) 113-117.

[17] Y. Degani, A. Patchornik, Cyanylation of sulfhydryl groups by 2-nitro-5-thiocyanobenzoic acid. High-yield modification and cleavage of peptides at cysteine residues, *Biochemistry* 13 (1974) 1-11.

[18] S. Nakagawa, Y. Tamakashi, Y. Ishibashi, M. Kawase, S. Taketomi, O. Nishimura, T. Fukuda, Production of hPTH(1-34) via a recombinant DNA technique, *Pept. Chem.* 31 (1993) 5-8.

[19] S. Nakagawa, Y. Tamakashi, Y. Ishibashi, M. Kawase, S. Taketomi, O. Nishimura, T. Fukuda, Production of human PTH(1-34) via a recombinant DNA technique, *Biochem. Biophys. Res. Commun.* 200 (1994) 1735-1741.

[20] W. Wu, W. Huang, J. Qi, Y. T. Chou, E. Torng, J. T. Watson, 'Signature sets', minimal fragment sets for identifying protein disulfide structures with cyanylation-based mass mapping methodology, *J. Proteome Res.* 3 (2004) 770-777.

[21] ASTM D 1193-99e1, "Standard Specification for Reagent Water", ASTM International.

[22] M. Cadene, B. T. Chait, A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins, *Anal. Chem.* 72 (2000) 5655-5658.

[23] R. K. Scopes, Measurement of protein by spectrophotometry at 205 nm, *Anal. Biochem.* 59 (1974) 277-282.

[24] J. R. Whitaker, P. E. Granum, An absolute method for protein determination based on difference in absorbance at 235 and 280 nm, *Anal. Biochem.* 109 (1980) 156-159.

[25] M. Larhed, C. Moberg, A. Hallberg, Microwave-accelerated homogeneous catalysis in organic chemistry, *Acc. Chem. Res.* 35 (2002) 717-727.

[26] N. S. Wilson, G. P. Roth, Recent trends in microwave-assisted synthesis, *Curr. Opin. Drug Discov. Devel.* 5 (2002) 620-629.

[27] S. T. Chen, S. H. Chiou, Y. H. Chu, K. T. Wang, Rapid hydrolysis of proteins and peptides by means of microwave technology and its application to amino acid analysis, *Int. J. Pept. Protein Res.* 30 (1987) 572-576.

[28] B. W. Li, Comparison of microwave oven and convection oven for acid hydrolysis of dietary fiber polysaccharides, *J. AOAC. Int.* 81 (1998) 1277-1280.

[29] V. J. Nesatyy, A. Dacanay, J. F. Kelly, N. W. Ross, Microwave-assisted protein staining: mass spectrometry compatible methods for rapid protein visualisation, *Rapid Commun. Mass Spectrom.* 16 (2002) 272-280.

[30] B. N. Pramanik, U. A. Mirza, Y. H. Ing, Y. H. Liu, P. L. Bartner, P. C. Weber, A. K. Bose, Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: a new approach to protein digestion in minutes, *Protein Sci.* 11 (2002) 2676-2687.

[31] H. Zhong, S. L. Marcus, L. Li, Microwave-assisted acid hydrolysis of proteins combined with liquid chromatography MALDI MS/MS for protein identification, *J. Am. Soc. Mass Spectrom.* 16 (2005) 471-481.

[32] S. S. Lin, C. H. Wu, M. C. Sun, C. M. Sun, Y. P. Ho, Microwave-assisted enzyme-catalyzed reactions in various solvent systems, *J. Am. Soc. Mass Spectrom.* 16 (2005) 581-588.

[33] E. Gross, J. L. Morell, The reaction of cyanogen bromide with S-methylcysteine: fragmentation of the peptide 14-29 of bovine pancreatic ribonuclease A, *Biochem. Biophys. Res. Commun.* 59 (1974) 1145-1150.

[34] G. T. Wernimont, W. Spendley, Use of Statistics to Develop and Evaluate Analytical Methods, AOAC Intl, Arlington, VA., 1985.

[35] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, How to measure and predict the molar absorption coefficient of a protein, *Protein Sci.* 4 (1995) 2411-2423.

[36] JMP IN software for statistical analysis, Version 5.1, SAS Institute Inc., Cary, NC., 2003.

[37] A. Miller, P. H. Solomon, Writing Reaction Mechanisms in Organic Chemistry, Second Edition, Academic Press, Orlando, FL., 1999.

[38] C. G. Swain, C. B. Scott, Concerted displacement reactions. X. Quantitative correlation of relative rates. Comparison of hydroxide ions with other nucleophilic reagents toward alkyl halides, esters, epoxides and acyl halides, *J. Am. Chem. Soc.* 75 (1953) 141-147.

CHAPTER 3. USE OF HOMOLOGOUS NUCLEOPHILES FOR CHEMICAL CLEAVAGE AND LABELING IN CYANYLATION-BASED DISULFIDE MASS MAPPING

1.- INTRODUCTION

1.1 Chemical Tagging Methodology for Quantitative Proteomics

Mass spectrometry has become the technology of choice for rapid identification of proteins in a biological mixture [1]. Mass spectrometry has several advantages over other analytical techniques; it has high sensitivity, accuracy and capacity. Tandem mass spectrometry provides an adequate tool to identify peptides and proteins through the two principal fragment ions that are produced; the so-called N- and C-terminal b- and y-type ions, respectively. High quality CID-MS/MS of enzymatic peptides show prominent b- and y-ion series. Database search algorithms can be used in conjunction with MS/MS to readily identify peptides and proteins. Computer-based identification of dozens to hundreds of proteins now can be attained in a single MS analysis [2].

There are several ways to face the more challenging problem of *de novo* sequencing (manual interpretation of peptide spectra for protein identification without a database). One is to reduce the complexity of peptide spectra by using stable-isotope labeling. Isotopic labeling can be performed by enzymatic

digestion of proteins in $^{16}O/^{18}O$ water, eliminating the need for peptide derivatization. Specifically, protein digestion by trypsin in the presence of 50% of $H_2{}^{18}O$ incorporates the ^{18}O atom selectively into the C-terminal carboxyl groups of peptides. Subsequent fragmentation by CID-MS/MS distinguishes the y-ion series by a characteristic 2-mass unit shift, which facilitates read-out of the sequence [3].

Other approaches for peptide sequencing and quantitative analysis applied in the genomics and proteomics fields involve protein-tag approaches. The MCAT (Mass-Coded Abundace Tagging) method relies on the selective and quantitative guanidation of the ε-amino group of the C-terminal lysine residues of tryptic peptides. At high pH, reaction with O-methylisourea selectively transforms lysine into homoarginine, which is 42 Da heavier than lysine. For peptide sequencing, the comparison of MS/MS spectra for each sister peptide pair permits the informative y-ion peaks to be readily discerned, allowing for systematic sequence determination (see Figure 56) [2]. Programs such as SEQUEST [4] can be useful, readily recognizing MCAT-derivatized peptides. For quantitative comparison of proteins present in samples derived from different cell states, one of the samples is treated with O-methylurea; then, the two samples are combined and analyzed by LC-MS. Full-scan MS spectra are recorded and the relative abundances of sister peptide species are determined by comparing the profiles of reconstructed single-ion chromatograms (see C in Figure 56).

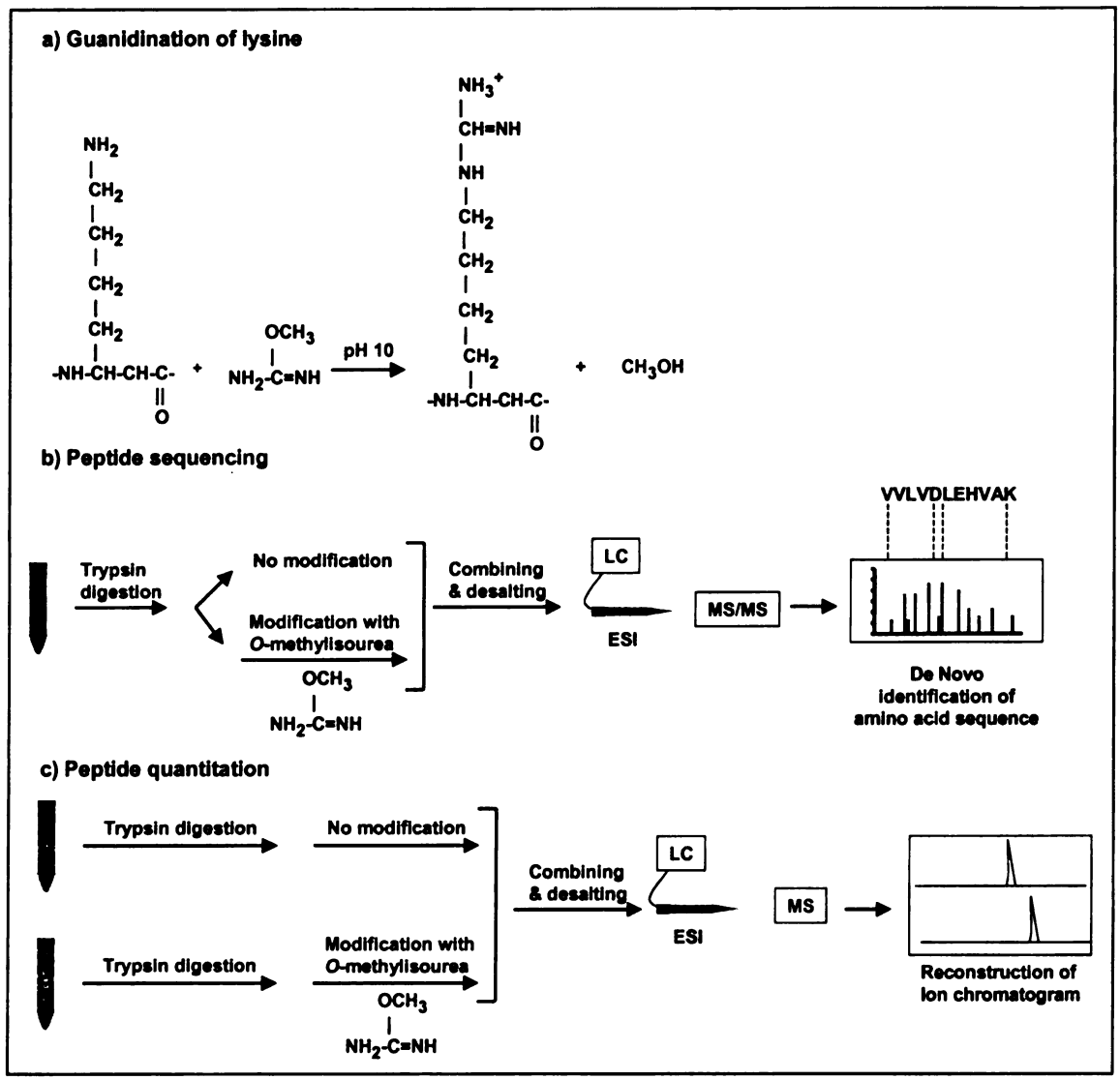


Figure 56. Overview of MCAT peptide sequencing and quantitation, from reference [2].

MCAT is based on the previously developed methodology for a quantitative analysis of complex protein mixtures developed by Gigy et al using Isotope-Coded Affinty Tags (ICATs) and tandem mass spectrometry. In this approach, the ICAT reagent consists of three functional elements: a chemically specific reactive group (e. g., with specificity toward thiol groups), an isotopically coded linker (to incorporate stable isotopes), and an affinity tag (e .g., biotin). An ICAT reagent with specificity to thiol groups is shown in Figure 57.

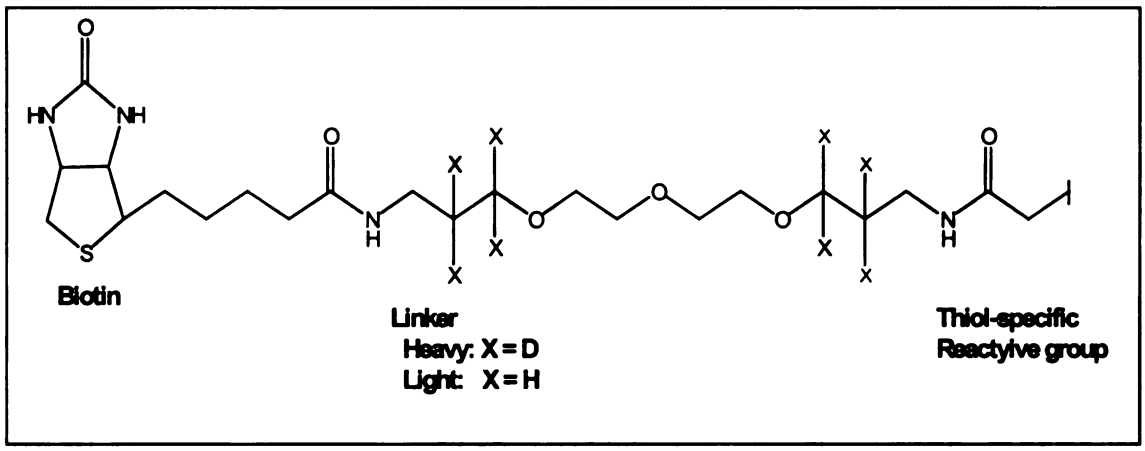


Figure 57. Structure of an ICAT reagent. The reagent exists in two forms, heavy (containing 8 deuteriums, X = D) and light (contains no deuteriums, X = H).

In the ICAT approach, proteins from cell state 1 are labeled using the light thiol-specific ICAT reagent (producing species d_0). Proteins coming from a cell state 2 are labeled with the heavy ICAT (producing species d_8). These protein mixtures are combined and proteolyzed to peptides. ICAT-labeled peptides are isolated using avidin affinity chromatography based on the biotin-avidin interaction; thus, only labeled peptides are isolated. Finally, the isolated peptides are separated and analyzed by μ-LC-MS/MS. A pair of ICAT-labeled peptides is chemically identical and is easily visualized because they essentially coelute, but there is an 8-Da mass difference measured in mass spectra for single charged species. The relative quantification is determined by the ratio of peptide pairs (d_o vs d_8), which represent peptides of the same protein in the two different states (See Figure 58) [5].

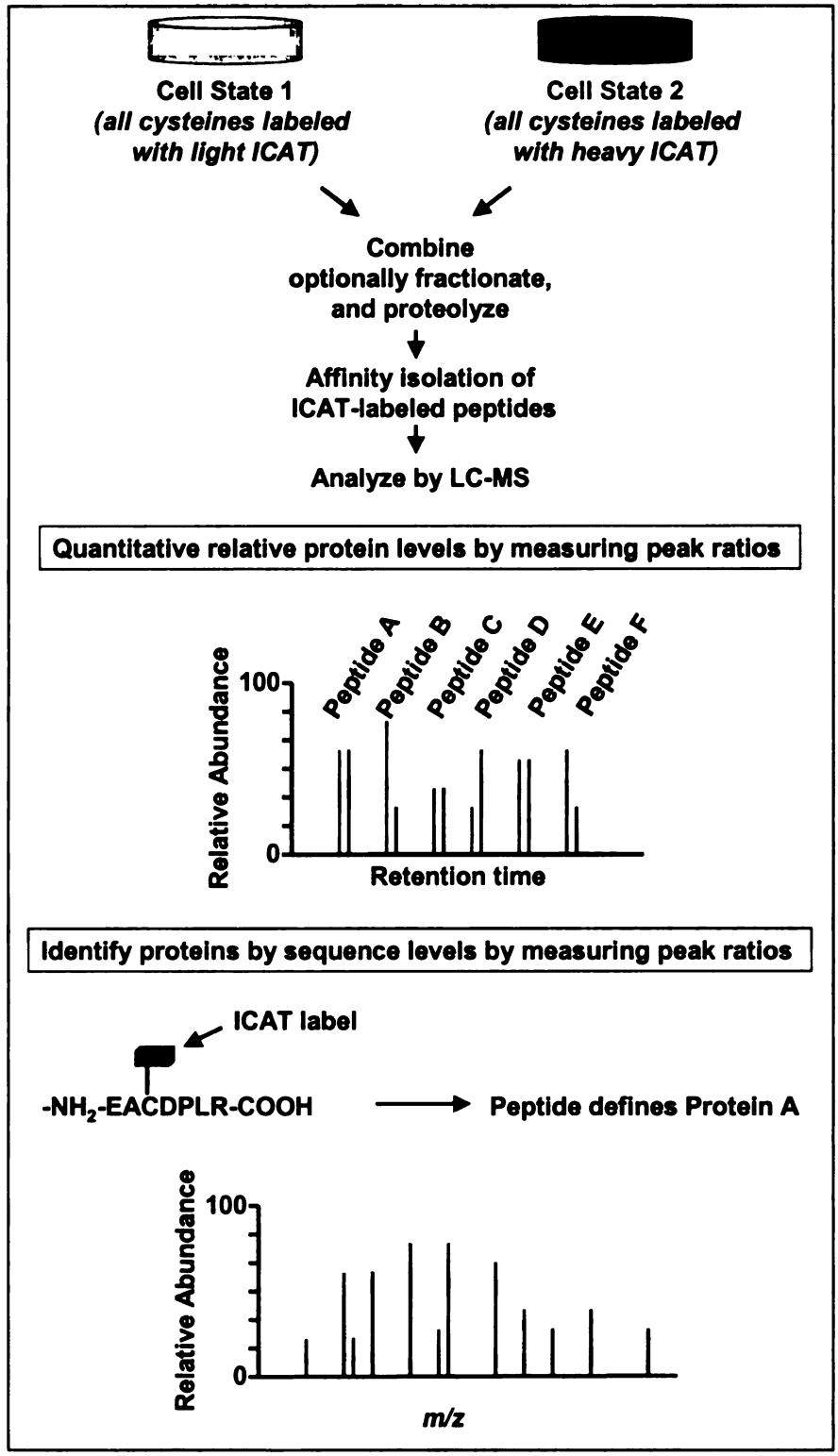


Figure 58. ICAT strategy for quantifying differential protein expression, based on thiol-specific ICAT reagents (adapted from reference [5]).

Wu et al proposed a methodology that uses brominated, structurally homologous reagents to quantify cysteinyl peptides. The two reagents used in this approach are 2-bromoacetamide (2-BA) and 2–bromopropinamide (2-BP). Even though the use of alkylating reagents could lead to problems of under- or over-modification, after optimization this strategy seems to be a good choice to analyze the relative concentration ratio of disulfide isomers [6].

1.2 Application of labeled methodologies in determination of disulfide bonds

As with the methods reviewed earlier, the use of methodologies where peptides are labeled are not common for the determination of disulfide linkages. However, the use of pepsin in solutions containing 50% of $H_2{}^{18}O$ has been proposed. After digestion, resultant disulfide-linked peptides have distinct isotope profiles compared to the same peptides with only ^{16}O in their terminal carboxylates. Thus, it is possible to recognize disulfide-linked peptide in digests and chromatographic fractions using these mass-specific markers, and to rationalize mass changes upon reduction in terms of half-cystinyl sequences of the protein of interest. The use of CID-MS/MS in this approach also is facilitated by the presence of ^{18}O in the peptide product of the enzymatic reaction [7, 8]. Even though some applications have been reported [9], this method is quite complicated experimentally, and its use is not widespread.

The cyanylation (CN)-based mass mapping methodology [10] for assignment of disulfide bond connectivities has proven useful even in proteins containing adjacent cysteines. As was explained in chapter two, the essence of the method is based on selective cyanylation of the sulfhydryl group of a cysteine residue, and subsequent nucleophilic attack to promote CN-induced cleavage on the N-terminal side of the modified cysteine. However, in some cases, the low yield or poor ionization of some cleavage fragments, makes data interpretation difficult because signals cannot be considered as "true signals" representing the cleavage fragments. Thus, it is possible to lose some valuable information that might be useful for a correct interpretation or for feeding algorithms that help with the assignment of the disulfide bonds [6, 11, 12]. We optimized experimental conditions using ammonia or methylamine to produce good yields of cleavage fragments in the context of the CN-based disulfide mass mapping methodology based on results from both UV-absorption HPLC chromatograms and MALDI-MS (see Chapter 2). The parallel use of two nucleophiles (e.g., ammonia and methyl amine) increases confidence in distinguishing cleavage fragments from chemical byproducts or noise.

During the nucleophilic cleavage step in CN-based disulfide mass mapping methodology, three cleavage fragments are produced: the N-terminal peptide (A), the middle iminothiazolidine (*itz*)-peptide (B), and the *itz*-C-terminal peptide (C). In the case of the N-terminal fragment (A) and the middle *itz*-peptide (B), the C-terminus of these peptides contains the nucleophile group (Nu). For

example, if the nucleophile is ammonia, the Nu group is -NH_2; in the case of methylamine, the Nu is –NH-CH_3; and for ethylamine, –NH-CH_2-CH_3, etc. Thus, when a binary mixture of two homologous nucleophiles is used, the A and B fragments will show a mass difference of 14 Da in MALDI-MS. On the other hand, the C-terminal fragment (C) will always show the same mass in spite the nucleophile used. This expected pattern improves confidence in detecting CN-induced cleavage products (see Figure 59).

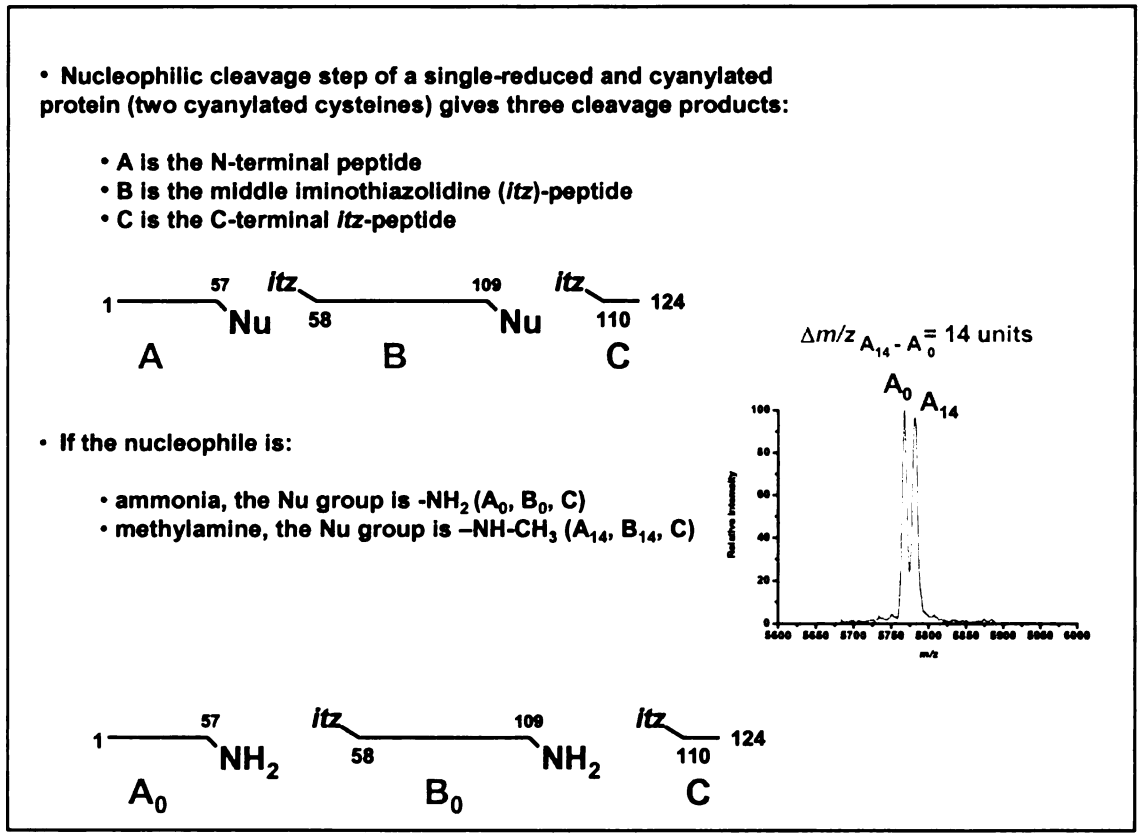


Figure 59. Impact of different nucleophiles on the cleavage reaction. Species produced using ammonia are designated with "$_0$" and cleavage fragments based on methylamine are designated with subscript "$_{14}$".

Proposed in this chapter is the development of a method to facilitate recognition of cleavage products using a binary mixture of homologous

nucleophiles during mass mapping of the disulfide structure of cystinyl proteins by means of the Partial Reduction/ Cyanylation/ Cleavage methodology (PRCC). Experimental results obtained after cleavage of the single-reduced/cyanylated isoforms of RNase-A using ammonia and methylamine are presented to show the feasibility of using a reagent containing both nucleophiles to produce pairs of labeled peptides that facilitate recognition of cleavage products A and B.

2.- MATERIAL AND METHODS

2.1 Chemicals

Guanidine hydrochloride was obtained from Boehringer Mannheim Biochemichals, (Indianapolis, IN); 1-cyano-4-dimethylaminopyridinium tetrafluoroborate (CDAP) and tris-(2-carboxyethyl)phosphine (TCEP) hydrochloride were purchased from Sigma (St. Louis, MO). Water grade I [13] or better was obtained from a Millipore system. Methylamine, was acquired as the hydrochloride salt from Aldrich. Ammonia was obtained from EM Science as an aqueous solution containing approximately 14.4 M ammonium hydroxide. Bovine pancreatic ribonuclease A type III-A (RNase-A) was purchased from Sigma (St. Louis, MO).

2.2 Mass spectrometry

MALDI mass spectra were acquired on a Voyager DE-STR time-of-flight (TOF) mass spectrometer (Perkin-Elmer Biosystems Inc.) equipped with a 337-nm nitrogen laser. Measurements were made in linear mode with the accelerating voltage set typically to 25,000 V with grid voltage at 95%, guide wire at 0.05%, and extraction delay time at 150 nsec. Time-of-flight to mass conversion was achieved by external and/or internal calibration using standards of bradykinin ($MH^+_{average}$ = 1061.22 Da), bovine pancreatic insulin ($MH^+_{average}$ = 5734.56 Da), horse skeletal myoglobin ($MH^+_{average}$ = 16,952 Da), and/or cytochrome c ($MH^+_{average}$ = 12,361.1 Da) obtained from Sigma (St. Louis, MO). Samples were prepared by spotting 0.5 μL of analyte + 0.5 μL of matrix solution onto a stainless steel sample plate with α-cyano-4-hydroxy cinnamic acid (from Sigma Chemical Co., St. Louis, MO) as the matrix (10 mg/mL) in the modified thin-layer method as described by Cadene and Chait [14] or by the conventional dried-droplet method. For some experiments, samples were desalted using Zip-Tips from Millipore (Billerica, MA).

2.3 Chromatography

HPLC was used for separation of mixtures using a Waters 2695 separation module coupled to a Waters 2487 Dual λ Absorbance detector at 214 nm. The column for purification of isoforms was a C4 (Vydac # 214TP54, 5-μm particle size, 300 Å pore, 4.6 x 250 mm) and for μ-HPLC separation of reaction mixtures, the column was a C18 (Vydac # 218TP5115, 5-μm particle size, 300-Å pore, 1.0 x 150 mm). HPLC fractions were collected manually and, if necessary, dried for further use.

2.4 Partial Reduction of RNase-A

The protein sample was placed in a vial where the partial reduction was carried out by adding TCEP in 0.1 M citrate buffer at pH 3.0 for the cysteine content and keeping a concentration of 4 nmol/μL (e.g., 40 nmol of TCEP was reacted with 10 nmol of RNase-A in a total volume of 10 μL) followed by incubation at room temperature for 15 min.

2.5 Cyanylation and isolation of single-reduced isoforms of RNase-A

A 20-fold molar excess of 0.1 M CDAP (in 0.1 M Citrate buffer at pH = 3.0) over the total cysteine content was added to the mixture of partially reduced isoforms for the cyanylation step at room temperature for 15 min. Single-reduced

and cyanylated species were purified using HPLC, and then dried. A 6M guanidine hydrochloride solution was used to reconstitute the sample in order to obtain solutions between 0.3 and 1 nmol/μL. Accurate concentration was determined, if necessary, by injecting an aliquot of this solution into a μ-HPLC system, and comparing the response to that of standards. After partial reduction and cyanylation, single-reduced/cyanylated isoforms of RNase-A were isolated using reverse phase-HPLC. Because the cyanylated isoforms elute after the intact protein (IP), we refer to them according to their elution order (IF1, IF2, IF3, IF4).

2.6 Cleavage of single-reduced and cyanylated RNase-A isoforms

The nucleophile solution was made using water grade I [13] or better. Preformed single-reduced cyanylated isoforms of RNase-A (~ 1 nmol in 1 μL of 6M guanidine) were treated with a given nucleophile or with a mixture of nucleophiles under optimized experimental conditions (2 °C, 10 min) of concentration, temperature, and reaction time.

2.7 Quenching of the cleavage reaction

After the cleavage reaction was allowed to proceed for the specified time, it was quenched by adding citric acid to establish pH = 3, followed by adding

0.1 M TCEP solution and incubating at 37 °C to completely reduce the remaining disulfide bonds.

3.- RESULTS AND DISCUSSION

3.1 Application of "external" labeling to locate CN-induced cleavage fragments

The simplest approach for the use of two nucleophiles to identify CN-induced cleavage fragments during the cyanylation-mass mapping methodology can be named as "external" labeling. It consists of the treatment of aliquots of a sample with different nucleophiles, for example, ammonia and methylamine in separate experiments. After the cleavage reaction takes place, it is quenched, and the protein is totally reduced prior to acquiring the mass spectra. Thus, in order to identify the cleavage fragments and consequently infer the disulfide linkage, the two individual spectra are examined for mass spectral peaks representing a specific mass shift, depending on the mass difference between the nucleophiles used. This methodology can be slightly modified by doing what follows: after the cleavage reaction and quenching, samples can be combined in order to obtain a single mass spectrum where the presence of pairs of labeled peptides is easily visualized. This is because there is a 14-Da mass difference in mass spectra for single-charged species; an appropriate algorithm could detect pairs of these labeled peptides easily.

Another variation to this approach is the *in silico* comparison of mass spectra of samples individually treated with the different nucleophiles. The computer would process and compare mass spectra that have been obtained individually. This approach is useful because pairs of labeled peaks can be visualized in a single spectrum even though they come from physically different experimental samples. One example of this approach is given in Figure 60, where MALDI mass spectra for one of the isoforms for α-lactalbumin were combined *in silico* in order to detect cleavage fragments. In this case, the single-cyanylated isoform after reduction of the disulfide bond between cysteine 73 and cysteine 91 was split into three aliquots, and they were treated with ammonia, methylamine, and ethylamine independently. The MALDI spectra were combined *in silico*, and cleavage fragments were identified. Figure 60 shows an expansion of the region between *m/z* 8200 and 8400 where mass spectral peaks corresponding to the cleavage fragment A for this isoform are labeled.

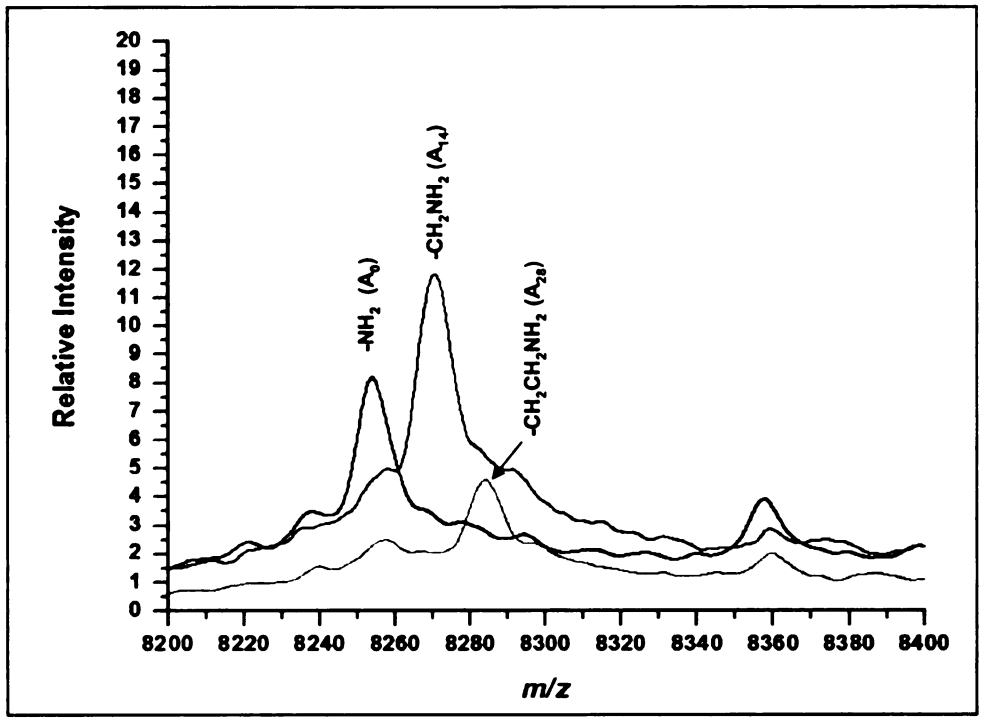


Figure 60. Expanded section of an overlay of three mass spectra processed *in silico* for one isoform of α-lactalbumin. Ao represents the cleavage fragment produced using ammonia as nucleophile, A_{14} when methylamine was the nucleophiles and A_{28} in the case of ethylamine.

We observe that the relative intensity of these peaks is very low (less than 10%), which would give us a high uncertainty in determining if these peaks are real signals from cleavage fragments if only one spectrum in a single experiment had been analyzed. However, in analyzing the combination of the three MALDI mass spectra, we are confident about assigning these signals to the A cleavage product coming from this particular isoform. In the spectrum, we observe that the intensity of the peaks is not similar. Ideally, peaks should have the same intensity, but this depends on the cleavage reaction and the MALDI-sample preparation. That is why we next looked for a single reagent that could produce

a pair of labeled cleavage fragments of similar peak intensity in a MALDI mass spectrum.

3.2 Optimization of the conditions for a combined ammonia-methylamine reagent for "internal" labeling

We define "internal" labeling for the CN-induced cleavage fragments as the use of a mixture of two nucleophiles as the cleavage reagent in order to produce a mixture of labeled cleavage fragments A and B in a single mass spectrum. The cleavage reaction takes place in one single sample and recognition of cleavage products A and B (see Figure 59) is especially facile when the concentrations of the homologous nucleophiles are adjusted to allow detection of equally intense peaks (relative intensity ratio ≡ rir = 1) as is shown in Figure 61.

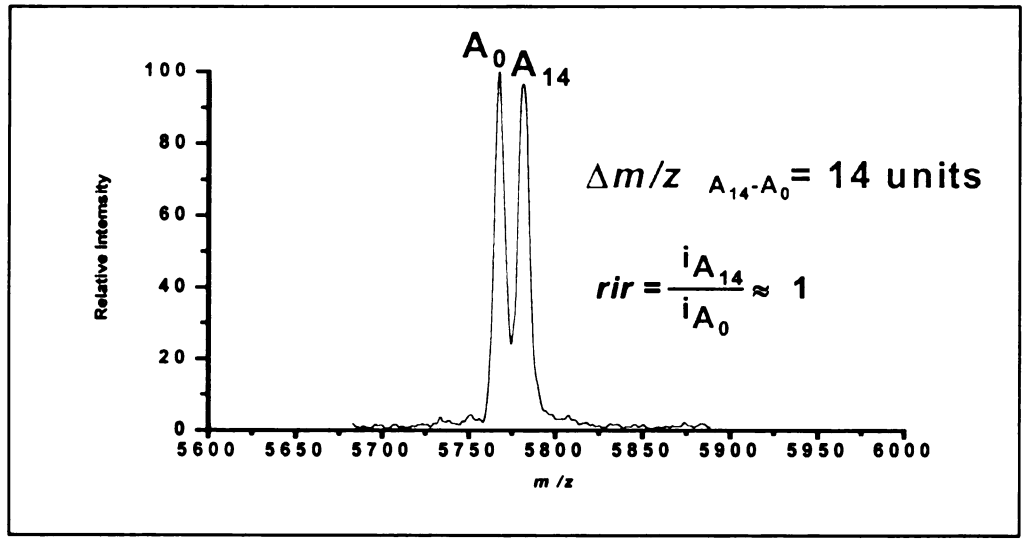


Figure 61. Ideal relative intensity ratio (rir) close to unity is represented for a MALDI signal for a pair of labeled cleavage fragments produced by treatment of the sample with a mixture of ammonia and methylamine.

In order to select the best nucleophile mixture based on the relative intensity ratio (rir) for the labeled cleavage products identified in MALDI mass spectra, singly reduced and cyanylated isoforms of RNase-A were chosen as model compounds. Taking into account the experimental conditions of temperature and reaction time (2°C, 10 min) previously evaluated for ammonia and methylamine, mixtures of these nucleophiles were evaluated for the production of cleavage products.

An aliquot of a stock solution of a single-reduced/cyanylated isoform was treated with a solution containing different methylamine/ammonia concentration ratios (see Table 6) under the optimized experimental conditions.

Table 6. Nucleophile concentrations for the reagent mixtures tested.

	$[NH_3]$ mol/L	$[MeNH_2]$ mol/L
a	5	0
b	5	0.14
c	5	0.21
d	5	0.28
e	5	0.35
f	5	0.42
g	5	0.49

Cleavage products A and B were identified in the MALDI mass spectra (see values in Table 7) and relative intensity ratios (rir) for the labeled cleavage products A and B were calculated from the obtained MALDI-MS spectra.

Table 7. Calculated *m/z* values for the protonated cleavage products using methylamine or ammonia as nucleophiles.

	A_0	A_{14}	B_0	B_{14}	C
IF1	7082.0	7096.0	788.9	802.9	5906.6
IF2	2704.9	2718.9	6546.5	6560.5	4526.1
IF3	4411.9	4426.0	6062.8	6076.8	3302.8
IF4	6351.2	6365.2	5766.5	5780.5	1659.8

Figure 62 shows the rir values for the labeled cleavage products of the different RNase A isoforms. Dashed lines show the ideal rir value (= 1). We observe that the rir for the labeled cleavage products reaches the value of 1 at slightly different nucleophile mixtures (depending on the isoform treated) because of the different accessibility of the cyanylated cysteines during the cleavage reaction.

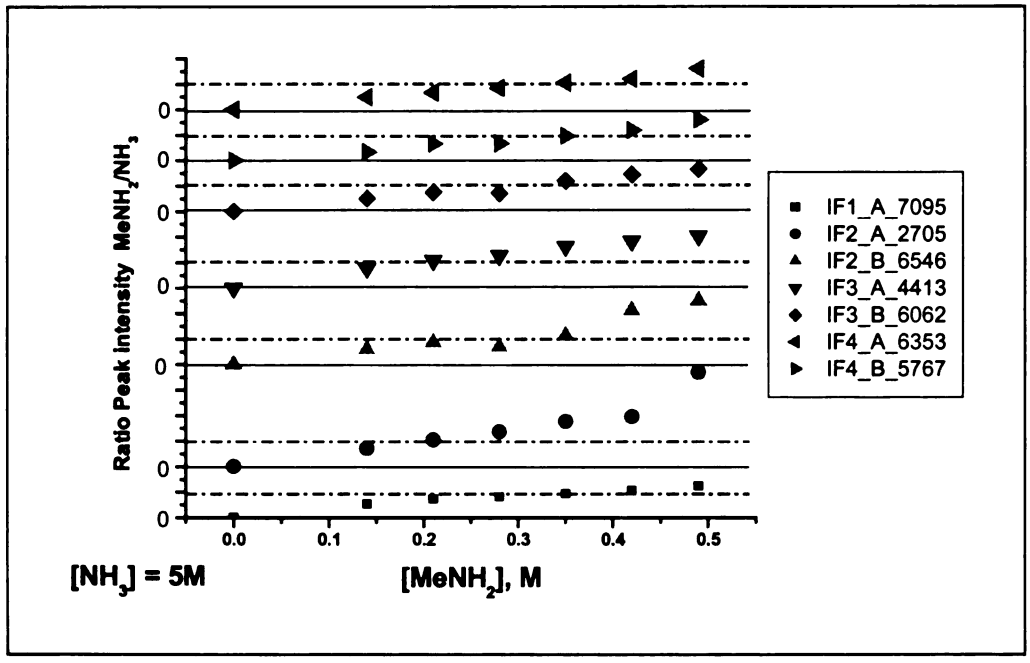


Figure 62. Relative intensity ratios for labeled cleavage products A and B for RNase A isoforms.

The graphed results indicate that a concentration ratio of 0.3M/5M for methylamine/ ammonia is a good compromise to obtain a rir of unity, which can can be corroborated by observing in detail the MALDI mass spectra under those experimental conditions.

Figure 63 and Figure 64 show the labeled A and B cleavage products for each of the RNase-A isoforms. From mass spectra, it is evident that the use of primary amines as cleavage reagents facilitates determination of the cysteine linkage in cystinyl proteins beacuse cleavage fragments are easily identified.

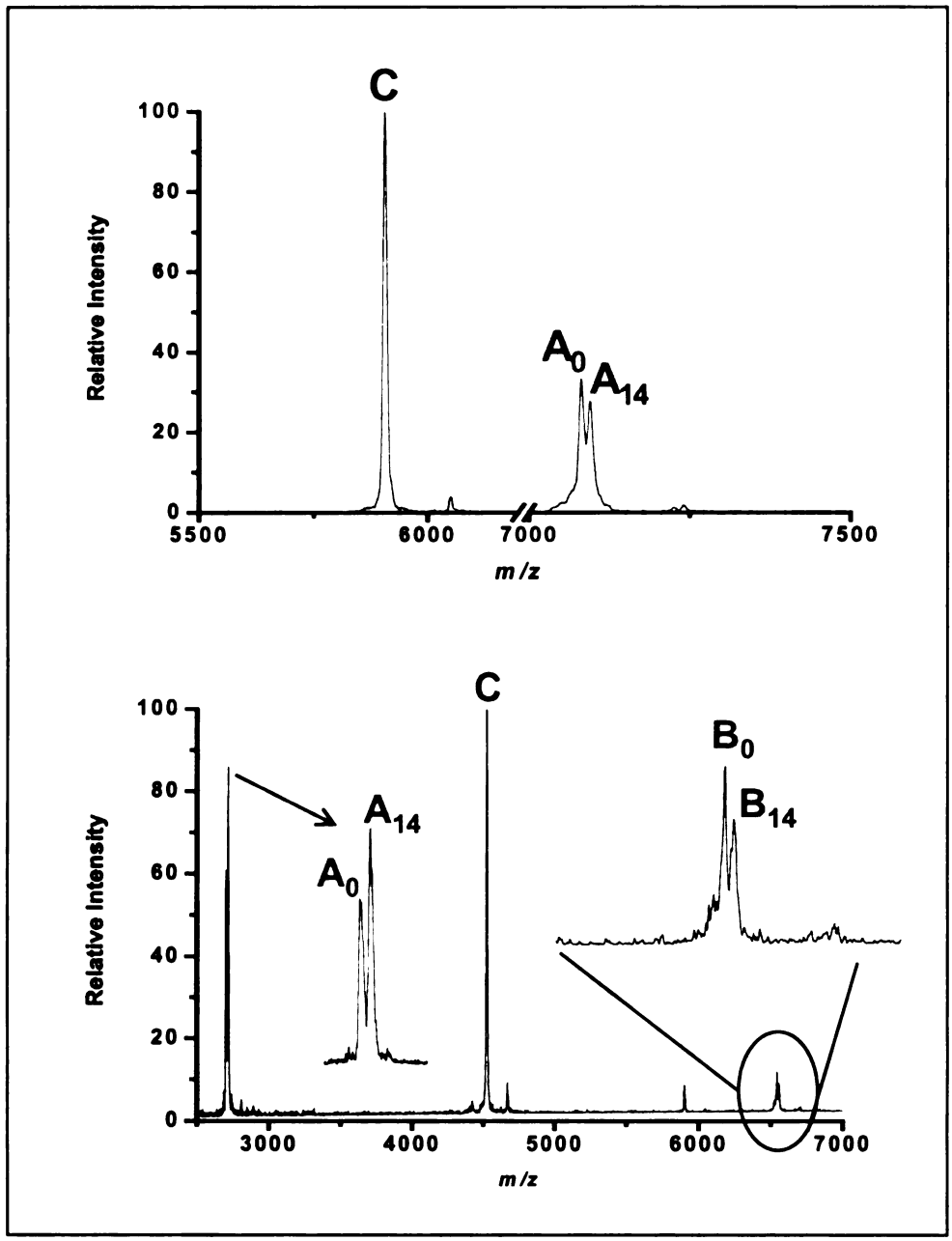


Figure 63. Top: Mass spectrum of labeled cleavage products for IF1. Bottom: Mass spectrum of labeled cleavage products for IF2.

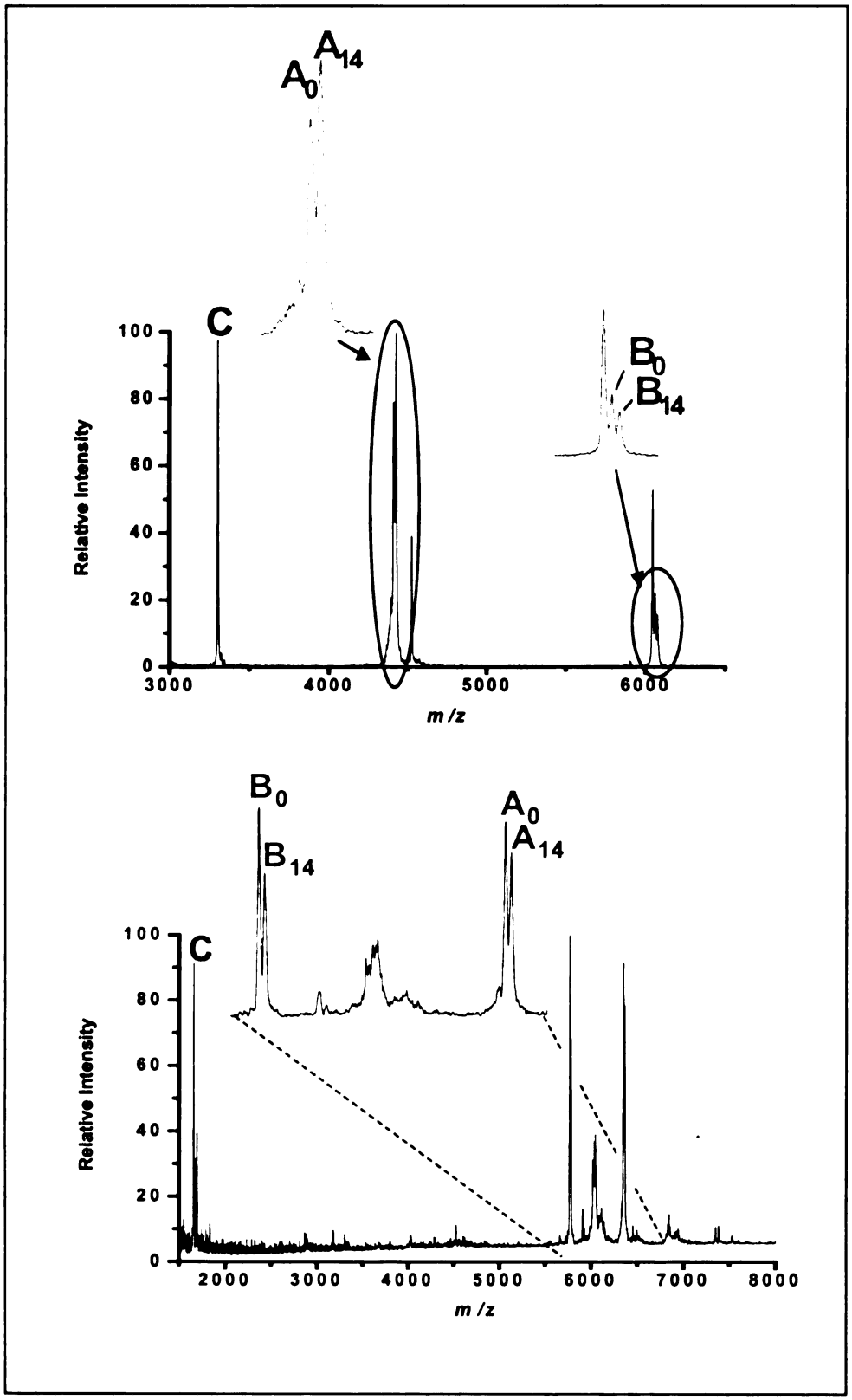


Figure 64. Top: Mass spectrum of labeled cleavage products for IF3. Bottom: Mass spectrum of labeled cleavage products for IF4.

4.- CONCLUSION

Mixtures of methylamine/ammonia as nucleophiles produce two different species for cleavage products A and B during the cleavage reaction of a cyanylated isoform.

Yields of the species produced during the cleavage reaction containing the nucleophile group $-NH_2$ or $-NH-CH_3$ depend mainly on the nucleophile concentration in the mixture even though there is an influence of the protein structure.

Methylamine is a better nucleophile than ammonia.

A concentration ratio for methylamine/ammonia of 0.3M/5M represents a good compromise to obtain a mixture of cleavage products A and B with MALDI mass spectra showing an intensity ratio equal to unity for species containing nucleophile groups $-NH_2$ and $-NH-CH_3$.

5.- REFERENCES

[1] A. Pandey, M. Mann, Proteomics to study genes and genomes, *Nature* 405 (2000) 837-846.

[2] G. Cagney, A. Emili, De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging, *Nat. Biotechnol.* 20 (2002) 163-170.

[3] A. Shevchenko, I. Chernushevich, W. Ens, K. G. Standing, B. Thomson, M. Wilm, M. Mann, 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 (1997) 1015-1024.

[4] J. K. Eng, A. L. McCormack, J. R. Yates, III, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database, *J. Am. Soc. Mass Spectrom.* 5 (1994) 976-989.

[5] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags, *Nat. Biotechnol.* 17 (1999) 994-999.

[6] W. Wu. *Applications and computational considerations of the cyanylation-based disulfide mapping methodology for cystinyl proteins*, PhD Thesis, Michigan State University, East Lansing, MI, December 2003.

[7] J. J. Gorman, T. P. Wallis, J. J. Pitt, Protein disulfide bond determination by mass spectrometry, *Mass Spectrom. Rev.* 21 (2002) 183-216.

[8] T. P. Wallis, J. J. Pitt, J. J. Gorman, Identification of disulfide-linked peptides by isotope profiles produced by peptic digestion of proteins in 50% (18)O water, *Protein Sci.* 10 (2001) 2251-2271.

[9] T. P. Wallis, C. Y. Huang, S. B. Nimkar, P. R. Young, J. J. Gorman, Determination of the disulfide bond arrangement of dengue virus NS1 protein, *J. Biol. Chem.* 279 (2004) 20729-20741.

[10] J. Wu, J. T. Watson, A novel methodology for assignment of disulfide bond pairings in proteins, *Protein Sci.* 6 (1997) 391-398.

[11] J. Qi, W. Wu, C. R. Borges, D. Hang, M. Rupp, E. Torng, J. T. Watson, Automated data interpretation based on the concept of "negative signature mass" for mass-mapping disulfide structures of cystinyl proteins, *J. Am. Soc. Mass Spectrom.* 14 (2003) 1032-1038.

[12] W. Wu, W. Huang, J. Qi, Y. T. Chou, E. Torng, J. T. Watson, 'Signature sets', minimal fragment sets for identifying protein disulfide structures with cyanylation-based mass mapping methodology, *J. Proteome Res.* 3 (2004) 770-777.

[13] ASTM D 1193-99e1, "Standard Specification for Reagent Water", ASTM International.

[14] M. Cadene, B. T. Chait, A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins, *Anal. Chem.* 72 (2000) 5655-5658.

CHAPTER 4. DEVELOPMENT OF AN ON-LINE SYSTEM FOR THE CYANYLATION-BASED DISULFIDE MASS MAPPING

1.- INTRODUCTION

1.1 On-line methods in analytical chemistry and proteomics

During many years, before and during the genomics and proteomics era, an increasing number of publications demonstrated the potential of column-switching techniques for the determination of drugs in biological and environmental samples using high performance liquid chromatography (HPLC). The use of these systems greatly facilitated the sample preparation step, providing better precision and accuracy in quantitative analysis and the possibility of time reduction in analysis [1-7]. In the simplest setup, a small column (trap) is located between the injection port and the analytical-HPLC column by means of a switching valve as illustrated in Figure 65. Two positions are available for this system. For the first position, the sample is loaded into the trap and the analytical column is conditioned and equilibrated at the same time. In the trap, the analytes are pre-concentrated, cleaned up and, if necessary, derivatized. Different washing steps with different solvents are carried out, if necessary, in order to remove impurities or undesirable analytes. However, one should be careful during this process to avoid eluting the analytes of interest during those steps. After analytes are ready to be eluted, the system is changed to position 2,

and the solvent gradient is started in order to desorb analytes from the trap, carrying out the separation by means of the analytical column (see Figure 65). Traps are filled with the appropriate stationary phase (resin) according to the analysis that is carried out. The most commonly used resins are reverse phase and ion exchange particles; however, some recent research suggested that the use of immunosorbents is a good choice for analysis of small molecules [8].

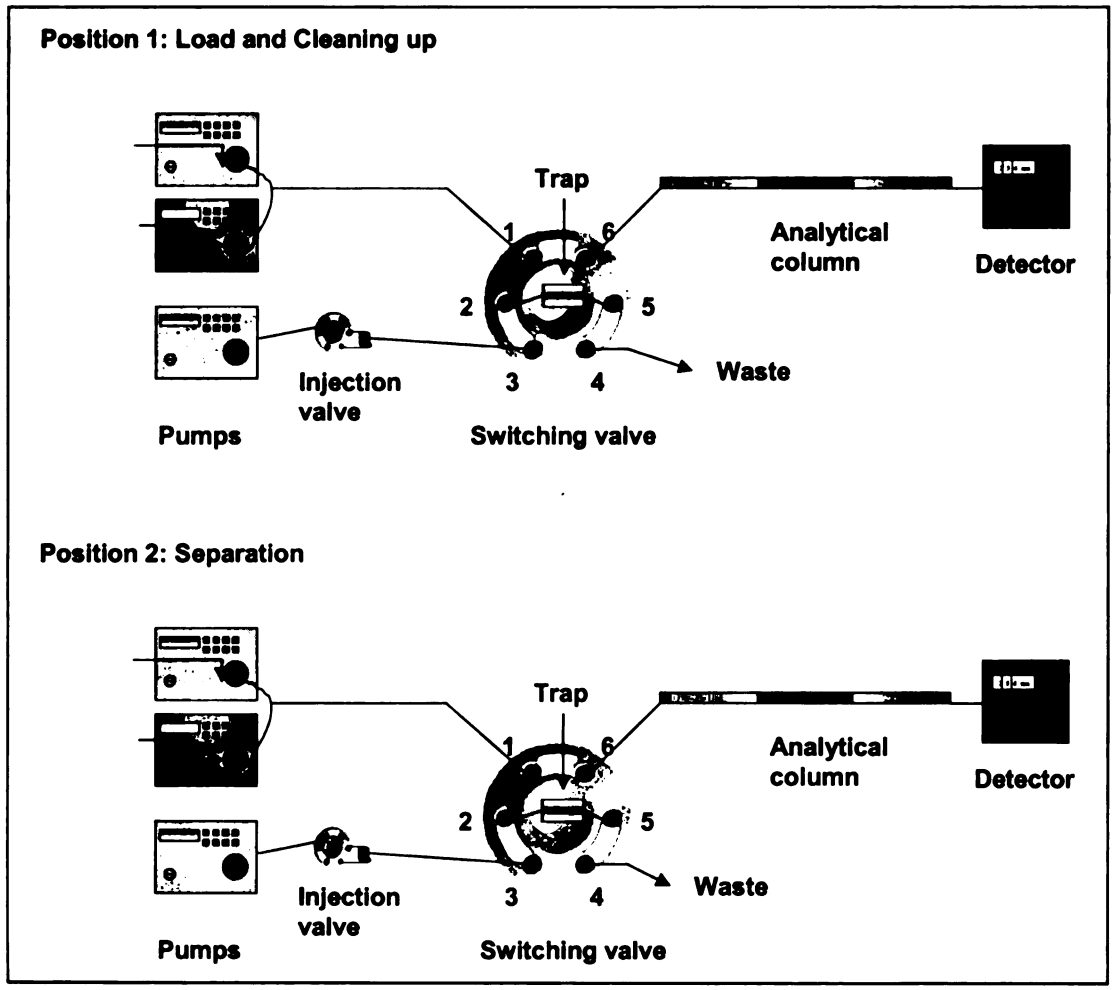


Figure 65. Simple on-line system for cleaning up and concentrating of samples. The system has two positions; read the text for details.

Automated HPLC systems were also used before the advent of biological mass spectrometry; anion-exchange columns were combined in tandem with reversed-phase columns for peptide mapping of very large proteins and the analysis of extremely complex peptide mixtures [9, 10]. Numerous reports exist concerning the use of tandem columns in liquid chromatography to address specific separation problems using column switching arrays or multidimensional chromatographic systems. In two-dimensional (2D) HPLC (2D-HPLC), the first column is filled with a stationary phase and it is operated under specific gradient conditions and separation mechanism, as the first dimension separation. Effluent from this first column is delivered into a second column for the second separation using a different mechanism and a different gradient program from the first separation (second dimension separation) [11-16]. The differences among several 2D-HPLC are based on the type of stationary phases used (separation mechanism) and/or the way columns are connected by means of switching and injection valves. An on-line two-dimensional HPLC-system is shown in Figure 66. This on-line system includes a preparation step with a silica-based restricted access material (RAM) with ion exchange functionalities. The separation is carried out in two dimensions by means of an ion exchange (IEX) column (first dimension) and four reverse phase (RP) columns (second dimension separation) that are used alternately as the fractions are eluted from the first separation. In this system, final fractions are collected for subsequent analysis.

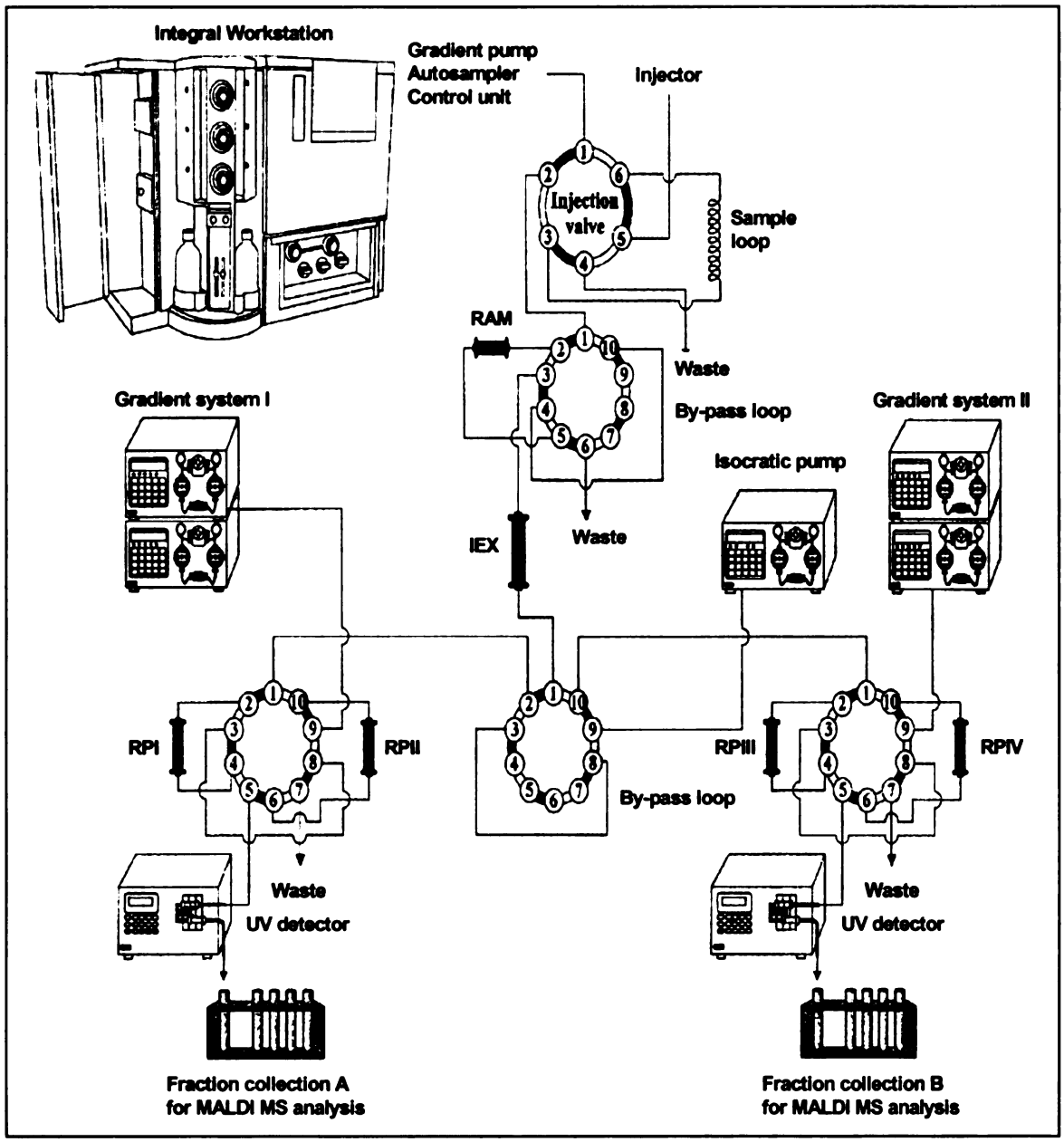


Figure 66. Schematic representation of an on-line system with sample preparation and two-dimensional HPLC. From reference [14].

Another interesting and quite novel strategy in liquid chromatography (LC) is called "thermally tuned tandem column (T^3C)". T^3C can be considered to be an optimal combination of adjusting selectivity by varying stationary phase type and temperature [17-19].

In the genomics and proteomics fields, switching and multi-dimensional separation alternatives have been used to couple separations of enzymatic digestions with mass spectrometry [12, 13, 20]. The necessity of working with small sample amounts have provoked the development of scaled down separation systems in order to detect peptide and protein quantities in the range of pico-, femto- and atto-moles. The use of membranes to immobilize proteins has also been developed for on-line systems applied in proteins [21-24]. The use of on-line digestions for proteins coupled on-line to mass spectrometry analysis is becoming more common, thereby improving the sensitivity of LC-hydrogen exchange-mass spectrometry experiments [24-28].

1.2 On-line method for the Partial Reduction-Cyanylation-CN Induced Cleavage (PRCC) Mass Mapping of Cysteinyl Proteins

Disulfide linkages play an essential role in protein tertiary structures. Changes in the disulfide bond structure of a protein may result in important changes of activity and function. The Partial Reduction-Cyanylation induced Chemical Cleavage for mass mapping of cysteinyl proteins is a methodology that

is used as a batch procedure, called the 'in-vial' approach, in which sequential reactions are carried out in a small vial [29]. During these steps, it is necessary to remove reagents and salts that have been added for one reaction before reagents can be added for the subsequent reaction. This chapter describes two different systems that were analyzed as possible analytical instrumentation arrangements for the PRCC method. In the first approach, only one small column, filled with a resin, was used for binding the protein and to carry out all chemical steps of the PRCC methodology. In the second approach, the on-line PRCC hardware consists of three components connected in series: T1-R-T2; a reactor (R) and two traps (T1, T2) containing polymeric resin. In this PRCC on-line system, the protein is loaded and subsequently chemically degraded in a controlled manner to give structurally diagnostic cleavage products, which are collected and analyzed using mass spectrometry (MS) in order to establish the cysteine linkage. The sample is loaded onto T1, where it is partially reduced and cyanylated. Then, cyanylated species are transferred to R where cleavage occurs in a continuous flow step. Species from R are captured in T2, and after total reduction, they are eluted for analysis by MS. This second instrumental configuration is the one recommended in the conclusions after a complete evaluation of the performance system, where the overall analysis time is reduced and the sample manipulation is minimized to help minimize sample loss. During development of the on-line PRCC system, the integrity of each physical or chemical step was tested with standard quantities of model compounds, primarily using ribonuclease A (RNase-A) and bradykinin (BDK). A comparison of the

general steps that are necessary for in-vial and on-line approaches is shown in Figure 67.

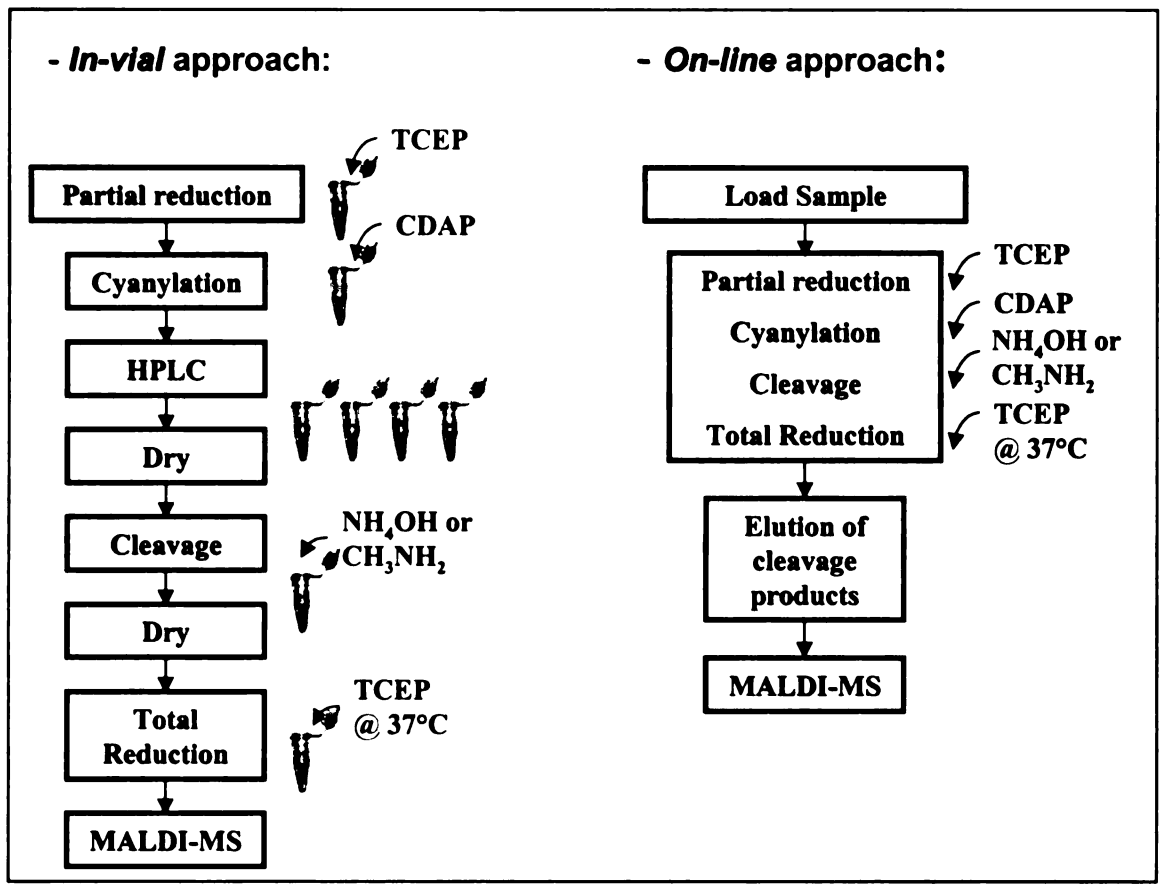


Figure 67. Comparison of the in-vial and on-line steps for the Partial Reduction Cyanylation and Cleavage methodology.

2.- MATERIAL AND METHODS

2.1 Chemicals and Enzymes

Guanidine hydrochloride was obtained from Boehringer Mannheim Biochemichals, (Indianapolis, IN); 1-cyano-4-dimethylaminopyridinium

tetrafluoroborate (CDAP) and tris-(2-carboxyethyl)phosphine (TCEP) hydrochloride were purchased from Sigma Chemical Co. (St. Louis, MO).

Water grade I [30] or better was obtained from a Millipore system.

Methylamine was acquired as hydrochloride salt from Aldrich. Ammonia was obtained from EM Science as an aqueous solution containing approximately 14.4 M ammonium hydroxide. Bovine pancreatic ribonuclease A type III-A (RNase-A, 8 cysteines, 4 disulfide bridges, 124 amino acid residues) and bradykinin (BDK, NO cysteines, 9 amino acid residues) were purchased from Sigma (St. Louis, MO).

2.2 Columns and reactor for the cleavage reaction

Empty microbore-HPLC columns (1.0 mm ID x 2 cm; V ~ 15.7 μL) made of PEEK©/stainless steel and narrowbore short-HPLC columns (2.0 mm ID x 2 cm; V ~ 62.8 μL), made of stainless steel from Upchurch (Oak Harbor, WA), were used to prepare lab-made traps (T) with different types of stationary phases. An empty analytical column was used as reactor (R) for the cleavage reaction (4.6 mm ID x 25 cm; V ~ 4.15 cm^3).

2.3 Stationary phases

PRP-3 polymeric reversed phase (pore diameter = 300 Å) was purchased from Hamilton (Reno, NV). Poros R2 was obtained from PerSeptive Biosystems (Framingham, MA); it is a polymeric stationary phase developed for perfusion chromatography. C4 and C18 silica-based stationary phases were used from available resins in the laboratory. A polymeric cartridge from Michrom was used also.

2.4 Preparation of traps with stationary phases

Empty traps were filled using a setup similar to the one shown in Figure 68. A slurry (usually 50:50 v/v methanol/water) containing the selected resin was deposited in the funnel and the vacuum carried the resin down the column (trap).

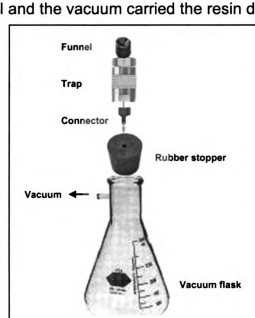


Figure 68. Lab-made setup to fill traps with stationary phase.

2.5 Mass spectrometry

MALDI mass spectra were acquired on a Voyager DE-STR time-of-flight (TOF) mass spectrometer (Perkin-Elmer Biosystems Inc.) equipped with a 337-nm nitrogen laser. Measurements were made in linear mode with the accelerating voltage set typically to 25,000 V with grid voltage at 95%, guide wire at 0.05%, and extraction delay time at 150 nsec. Time-of-flight to mass conversion was achieved by external and/or internal calibration using standards of bradykinin ($MH^+_{average}$ = 1061.22 Da), bovine pancreatic insulin ($MH^+_{average}$ = 5734.56 Da), horse skeletal myoglobin ($MH^+_{average}$ = 16,952 Da), and/or cytochrome c ($MH^+_{average}$ = 12,361.1 Da) obtained from Sigma Chemical Co. (St. Louis, MO). Samples were prepared by spotting 0.5 μL of analyte + 0.5 μL of matrix solution onto a stainless steel sample plate with α-cyano-4-hydroxy cinnamic acid (from Sigma Chemical Co., St. Louis, MO) as the matrix (10 mg/mL) in the modified thin-layer method as described by Cadene and Chait [31] or by the conventional dried-droplet method [32, 33].

2.6 Partial Reduction, Cyanylation and Isolation of Single-Reduced Isoforms of RNase-A

During evaluation of the online system for the PRCC methodology, it was necessary to work with single-reduced or single-reduced/cyanylated isoforms. Their preparation was according to the description in Chapter 2. Accurate concentration was determined, if necessary, by injecting an aliquot of standard solutions into μ-HPLC or narrowbore-HPLC systems and comparing the peak areas in the corresponding chromatograms versus areas in a calibration curve. Because the cyanylated isoforms for RNase-A elute after the intact protein (IP), we refer to them according to their elution order (IF1, IF2, IF3, IF4).

2.7 Chromatography

HPLC was used for separation of mixtures using a Waters 2695 separation module coupled to a Waters 2487 Dual λ Absorbance detector at 214 nm. The column for purification of isoforms contained a C4 stationary phase (Vydac # 214TP54, 5-μm particle size, 300 Å pore, 4.6 x 250 mm), the column for narrowbore-HPLC separation of reaction mixtures contained a C18 sationary phase (Vydac # 218TP3215, 3-μm particle size, 300-Å pore, 2.1 x 150 mm) and the column for μ-HPLC separation of reaction mixtures contained a slightly different C18 stationary phase (Vydac # 218TP5115, 5-μm particle size, 300-Å

pore, 1.0 x 150 mm). HPLC fractions were collected manually and, if necessary, dried for further use.

3.- RESULTS AND DISCUSSION

3.1 One-column approach for the on-line PRCC methodology

3.1a Proposed hardware for the one-column on-line system

This approach (see Figure 69) consisted of the use of only one trap (trap 1, T1) filled with a reverse stationary phase; it is connected on-line to a pumping system and a detector. The protein is injected into the system by means of the injection valve, loaded into T1, and bound to the stationary phase; reagents are added consecutively to carry out the chemical derivatization. Washing steps for removing of the chemical reagents are monitored through the UV-Vis detector. After total reduction (the last step), T1 can be connected to a HPLC system for partial separation of the cleavage products. Chromatographic fractions can be collected for a later analysis by MALDI-MS or can be analyzed by coupling a MS system online. In summary, the steps to carry out the PRCC methodology using this setup are:

1) Load sample on T1

2) Add TCEP for partial reduction

3) Washing step

4) Add CDAP for cyanylation

5) Washing step

6) Add Cleavage reagent

7) Washing step

8) Add TCEP for total reduction

9) Elute and collect cleavage products from trap using a HPLC coupled to T1.

10) Analysis using Mass Spectrometry (MALDI/TOF)

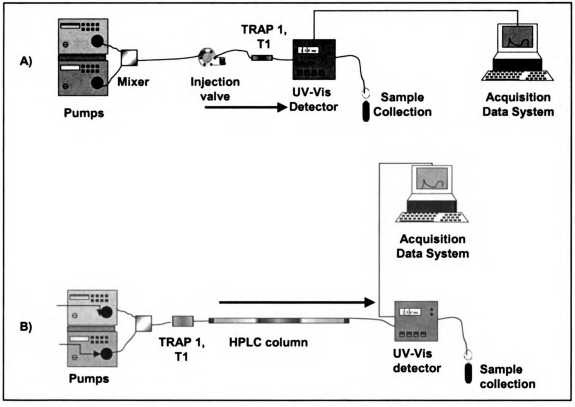


Figure 69. Single-trap approach for an on-line method for PRCC methodology. The instrumental setup in A) is used for the PRCC methodology (steps 1-8), after this, T1 is connected to a HPLC system (as in B) for elution and collection of cleavage products (step 9).

3.1b Evaluation of the one-column on-line approach

The first issue to be solved was the selection of the packing material for the trap 1; silica-based stationary phases dominate the field of separation science because of the high level of reproducibility, selectivity and sensitivity they can provide. However, these types of stationary phases are only hydrolytically stable over the pH range ~ 2-8.5 [27, 34, 35]. Thus, even though the use of C4 or C18 silica-based stationary phases had been shown to be suitable for the partial reduction and cyanylation of a cystinyl protein bound to the resin, it is necessary to select yet another stationary phase, because during the cleavage step, the pH is higher than 8.0. The inherent stability of PRP-3 columns (a cross-linked poly(styrene-divinyl benzene) polymeric stationary phase as illustrated in Figure 70) allows their use for any protein purification method in extreme pH-conditions from pH 1 to 13 (there is no silica to dissolve at pH higher than 8 or siloxane bonds in the C18 groups to hydrolyze at pH lower than 2) [36]. The PRP-3 support is pressure-stable up to 5,000 psi, the crosslinking prevents shrinkage or swelling when the mobile phase is changed. This kind of resin has been used successfully for recovering relatively non-polar analytes in on-line systems used in environmental chemistry [3, 37]. This PRP-3 resin is very helpful for separation of proteins that are unstable or have poor solubility at a pH below 8.0. While most proteins are best chromatographed at pH 2.0, it is good to know that they can be separated at elevated pH, if the need arises. Thus, PRP-3 columns allow processing of proteins at low and high pH extremes.

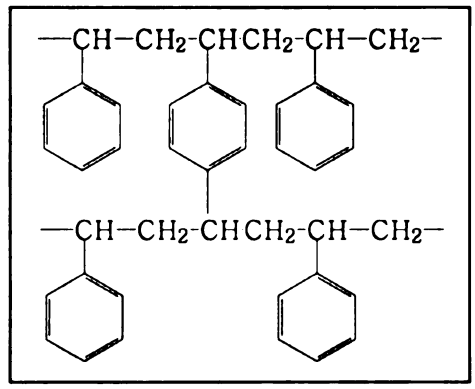
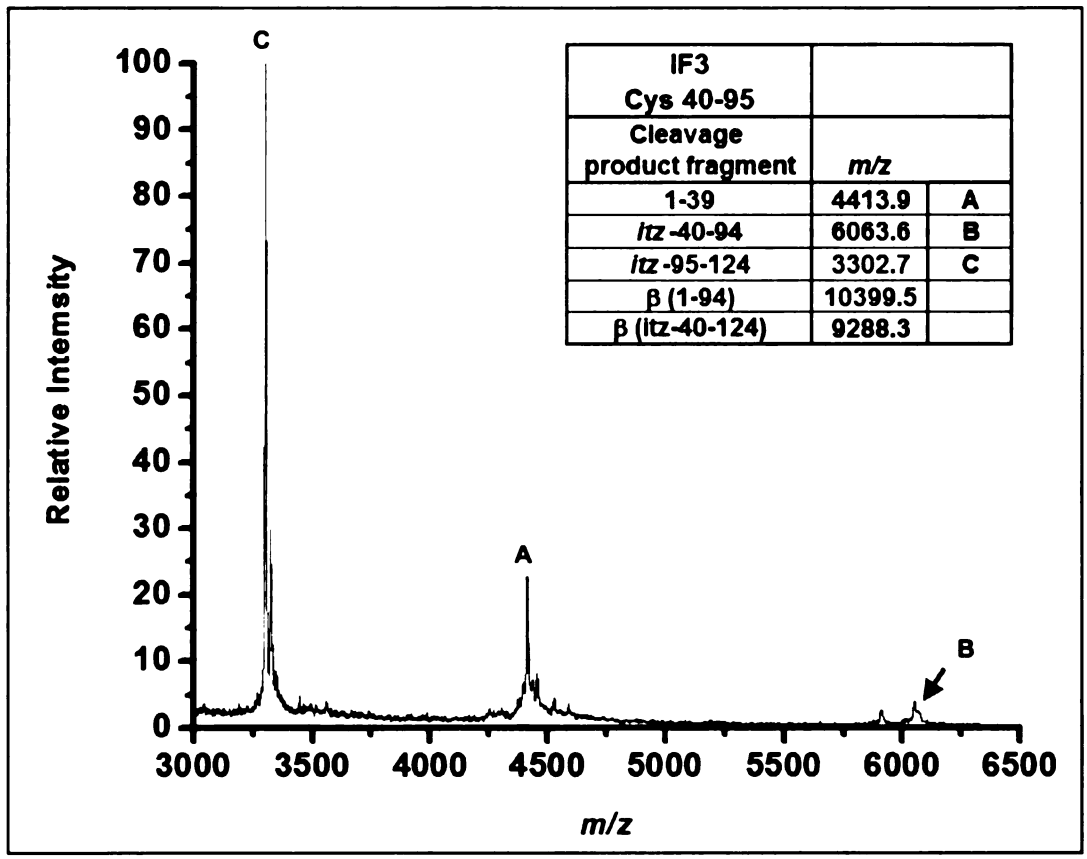


Figure 70. Chemical structure of PRP-3: poly(styrene-divinyl)benzene resin

According to the manufacturer, the 300-Å pores on the PRP-3 support help ensure good recovery (>90%) of the protein, and the highly inert poly(styrene-divinylbenzene) packing enhances protein recovery because there are no silanol groups on the support to cause irreversible protein adsorption [38].

Initial screening experiments showed that it was feasible to carry out partial reduction and cyanylation of the cystinyl protein (RNase-A) as it was bound to the PRP-3 resin, thus, it was decided also to evaluate the chemical cleavage reaction in the presence of the pRP-3 resin. Pre-formed single-reduced/cyanylated species of RNase-A (IF1... IF4) were loaded into T1; cleavage reagent (in this case, 1 M ammonia) was added to the trap and held there for the designated reaction time(s). After this, the excess cleavage reagent was washed off, reducing agent was added for total reduction, and finally the cleavage products were eluted from T1 and collected to be analyzed by MALDI-TOF MS. Results from screening experiments for the cleavage reaction showed that for IF1, IF2, and IF3, the cleavage reaction took place and cleavage

fragments were detected easily. The MALDI-MS spectrum for IF3 cleavage products as produced on-line under the described experimental conditions is shown in Figure 71. Mass spectral peaks for cleavage products are easily discernible.



IF3 Cys 40-95		
Cleavage product fragment	*m/z*	
1-39	4413.9	A
Itz-40-94	6063.6	B
Itz-95-124	3302.7	C
β (1-94)	10399.5	
β (Itz-40-124)	9288.3	

Figure 71. MALDI-MS spectrum for IF3 of RNase-A after being loaded into a small column packed (Trap1) with PRP-3 resin.

However, in the case of IF4, disappointing results were obtained. In Figure 72, the MALDI-MS spectrum is shown for on-line and in-vial approaches. It turned out that in using the on-column approach, two of the cleavage products could not be detected. Furthermore, none of the intact isoform could be detected either.

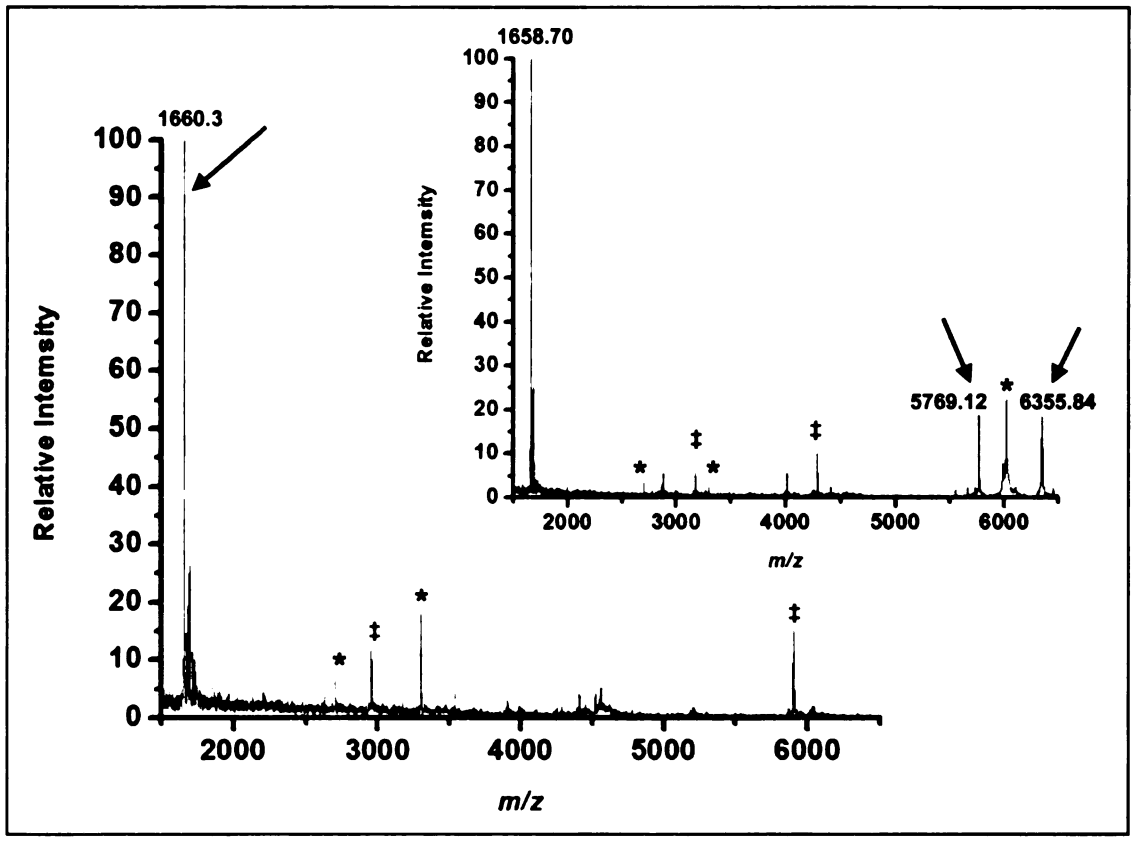


Figure 72. MALDI-MS spectrum for IF4 of RNase-A obtained using in-vial or on-column approaches. During the in-vial approach, two of the cleavage fragments (pointed out with arrows) are not observed. Peaks corresponding to i) cleavage fragments of minor cyanylated isoforms that coeluted with IF4 or β-elimination products are designated with (*) and ii) those some unknown side-reaction products are designated with ‡.

Table 8 shows a list of a series of experiments that were carried out in order to evaluate: 1) whether the cleavage products might be irreversibly adsorbed to the stationary phase in T1 and 2) other experimental conditions for the on-line cleavage reaction. For the first question, preformed cleavage-products coming from IF4 were loaded into trap 1. T1 was washed with water (0.1% TFA), and cleavage products retained in the trap were eluted using an ACN/water mixture. Eluate was analyzed by MALDI-MS. From MALDI-MS spectra, it was possible to identify all the cleavage products previously loaded

into T1. Thus, according to these results, there was not an irreversible binding of the IF4-cleavage products to the stationary phase.

Table 8. Checking list of experimental conditions tested for cleavage of IF4 on-column.

a) Evaluate Interactions Resin-Cleavage Products
✓ **Load and elution of pre-formed cleavage products (Residence time: 1 hr and overnight)**
b) Cleavage Reaction Variations
✗ **1 M NH_3 in 6 M Guanidine x 1 hr @ room temperature**
✗ **1 M NH_3 in 6 M Guanidine x 1 hr @ 37°C**
✗ **1 M NH_3 in 6 M Guanidine x 1 hr @ 50°C**
✗ **1 M NH_3 in 1 hr @ room temperature**
✗ **3.7 M NH_3 in 6 M Guanidine x 1 hr @ room temperature**
✗ **Concentrated NH_3 x 1 hr @ room temperature**
✗ **3.7 M NH_3 in 6 M Guanidine x 1 hr (using sonicator)**
✗ **1 M NH_3 in 6 M Guanidine x 1 hr @ room temperature (Flow rate aprox. 0.02 mL/min)**
✗ **1 M NH_3 in 6 M Guanidine in presence of a surfactant**

For the other experiments, no improvement was seen in the production of cleavage products from IF4. Analyzing the structure of RNase-A, using a Kyte & Doolittle plot for hydrophobicity of the protein [39], it is evident that cysteines that form the disulfide bond for IF4 are located in regions of high hydrophobicity in the protein (see Figure 73). Hydrophobic interactions retain the protein on the resin, and probably affect the way that the protein lies on the stationary phase; these interactions also might promote secondary reactions to compete with the

nucleophilic attack involved in the cleavage reaction and by consequence, diminish the yield of the expected cleavage fragments, in which case they may not detected in the mass spectra.

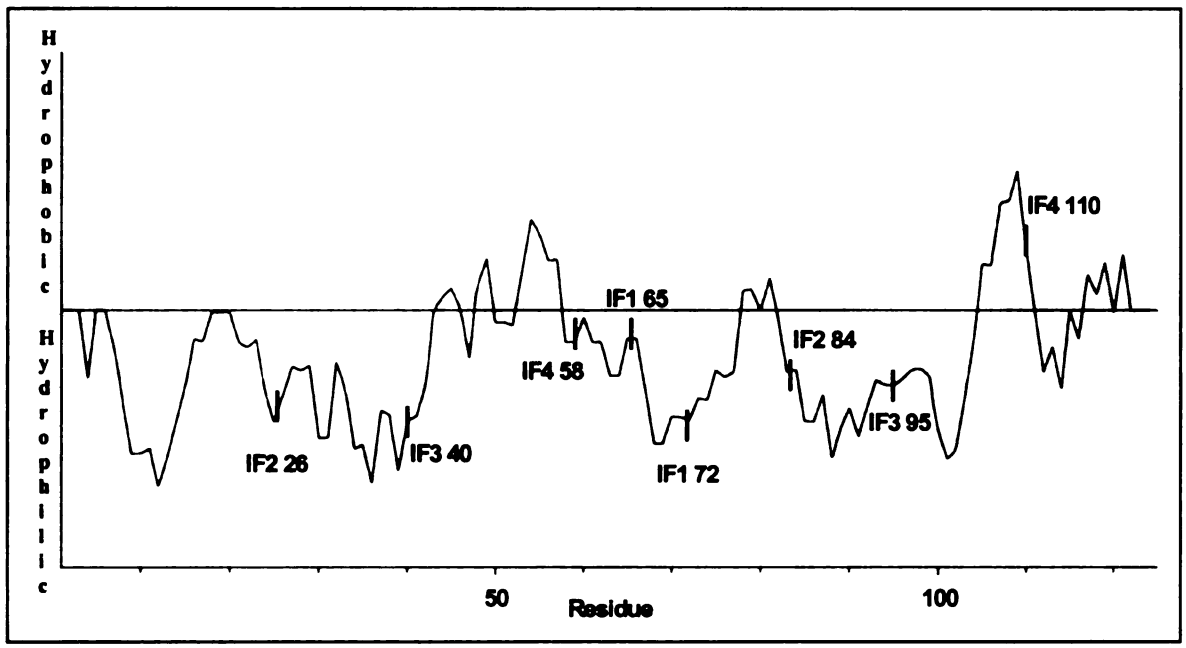


Figure 73. Kyte & Doolittle hydrophobicity plot for designated single reduced isoforms of RNase-A. The location of the cysteine contained in each isoform is indicated after the numeral identifying the isoform.

Thus, after the analysis of results obtained for the possible cleavage of the cystinyl protein adsorbed in a polymeric stationary showed this reaction was not always carried out efficiently, and it was necessary to think about a different approach for the PRCC methodology on-line. In this second approach the protein is desorbed from the trap, cleaved in solution, and trapped in a second trap. The next section describes and evaluates this alternative approach.

3.2 Two-column approach for on-line PRCC methodology

3.2a Proposed hardware for the two-column on-line system

This second approach consists of the use of two traps (T1, T2) for the on-line methodology. Partial reduction and cyanylation are carried out in T1 (see Figure 74). Cyanylated species are eluted from this trap and they are transported from T1 to T2 in a stream of the cleavage reagent through a reactor (R) while the cleavage reaction takes place. Formed cleavage products are captured by means of T2 where, after washing to remove the cleavage reagent, the remaining disulfide bonds are reduced by adding a reducing agent (TCEP in this case).

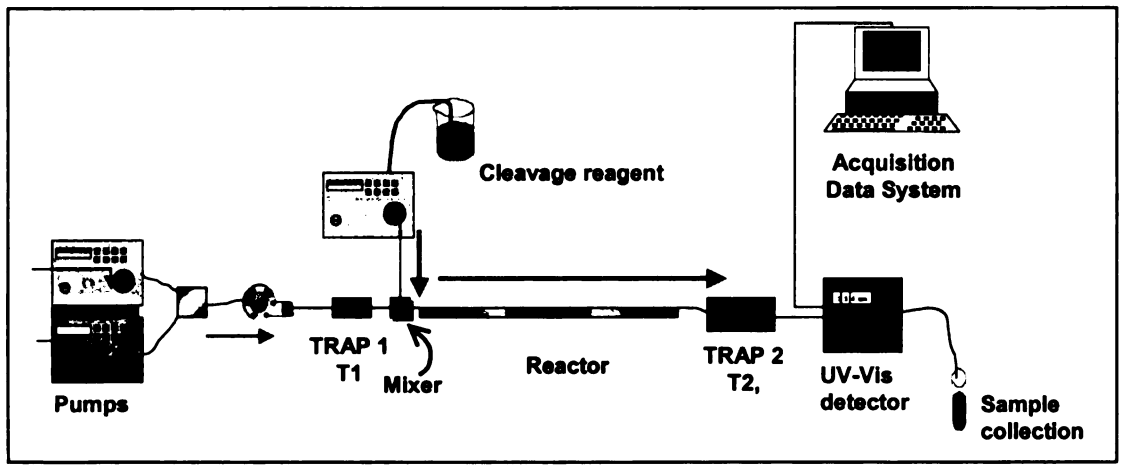


Figure 74. The two-trap and reactor approach for an on-line method for PRCC methodology.

Without taking into account the washing steps, the sequential steps necessary to complete the PRCC methodology using this setup are:

1) Load sample on Trap 1

2) Add TCEP to Trap 1 for partial reduction

3) Add CDAP to Trap 1 for cyanylation

4) Cleave the cyanylated isoforms during transfer from Trap 1 to Trap 2

5) Add TCEP for total reduction in Trap 2

6) Elute Cleavage Products from Trap 2

7) Analyze by MALDI-MS

In order to accomplish these steps, it was necessary to evaluate different parameters for the components involved in the setup: trap 1, the reactor, and trap 2. Parameters that were evaluated are summarized in Table 9.

Table 9. Evaluated parameters to test the feasibility of the two-column approach for developing an on-line system for the PRCC methodology.

TRAP 1	REACTOR	TRAP 2
• RESIN SELECTION • LOAD CAPACITY EFFICIENCY • FEASIBILITY OF CHEMICAL REACTIONS: PARTIAL REDUCTION AND CYANYLATION • DESORPTION EFFICIENCY	• CLEAVAGE REACTION EFFICIENCY • TRANSFER EFFICIENCY FROM TRAP 1 TO TRAP 2	• CAPTURE OF CLEAVAGE PRODUCTS

3.2b Trap 2 evaluation: Capture of cleavage products

Trap 2 was evaluated first because if T2 was not able to efficiently retain cleavage products under the cleavage experimental conditions, it would not have been possible to use the proposed setup in Figure 74. For T2, we used a narrowbore short-HPLC column filled with Poros R2 stationary phase. The Poros R2 stationary phase is a revolutionary product invented by Perseptive Biosystems for perfusion chromatography. This material consists of cross-linked poly(styrene-divinylbenzene) flow-through particles with a patented pore size distribution. This resin possesses large throughpores (6000-8000 Å) that allow analyte molecules to "perfuse" rapidly through the interior of the particles, as well as through very short "diffusive" pores (500-1500 Å) [27, 40].

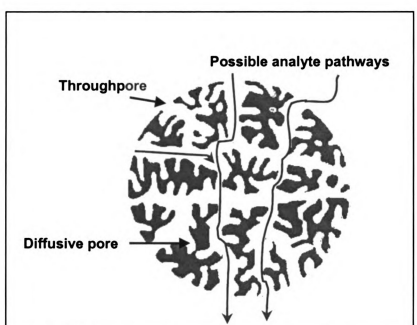


Figure 75. Diagram of a perfusion particle.

An additional property of Poros R2 is that it offers the lowest backpressure among the many types of resin in the marketing. Thus, knowing that Poros R2 is also a polymeric resin, it was decided to use this type of stationary phase for T2. RNase-A and BDK were chosen as model proteins. BDK was chosen because, after the chromatographic analysis of all cleavage products coming from the cyanylated RNase-A isoforms, it was noticed that BDK elutes approximately at the same time as the least retained cleavage products from RNase-A. Flow rate and percentage of acetonitrile that allow an effective trapping of cleavage products in trap 2 were evaluated. T2 was connected according to the diagram shown in Figure 76. Parameters were evaluated according to the methodology described next:

1) The cleavage reaction uses ammonia at a high pH, so a mixture of acetonitrile/ammonia solution was adjusted to the desired proportion of reagents and it was pumped through T2 following the "load" pathway that is shown in Figure 76.

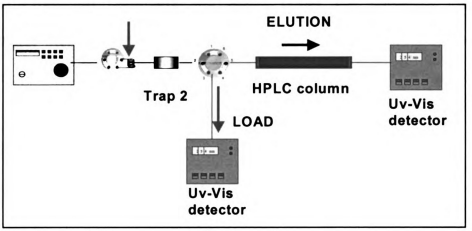


Figure 76. Scheme of the instrumental setup used to evaluate the capture of cleavage products by Trap 2.

2) A volume of 100 mL of sample mixture (1 nmol of RNase-A and 5 nmol of BDK in a solution of 6 M guanidine) was injected through the injection valve while the solvent mixture was continuously pumped (continuous flow). Adsorption of the proteins was followed by means of an UV detector. On average, the loading time was 15 minutes.

3) Trap 2 was washed with a solution of H_2O to remove the excess of the ACN/NH_3, and it was connected on-line to an analytical column through the "elution" pathway shown in Figure 76.

4) A gradient program of acetonitrile/water was carried out to elute RNase-A and BDK adsorbed onto Trap 2.

5) From the chromatogram, areas under the curve of the peaks representing the model proteins were calculated and compared with standard solutions in order to determine the amount of proteins adsorbed in Trap 2.

Thus, following the steps described above, two different sets of experiments were carried out:

1) Evaluation of the maximum content of ACN: RNase-A and BDK were adsorbed into Trap 2 using a flow rate of 100 μL/min using a solvent mixture of 0, 5, and 10% of ACN in 1.4 M ammonia. From results obtained here, it was decided that trapping of the model protein from a solution containing more than 5% ACN was inefficient, and by consequence it could be inferred that, for cleavage products, a concentration of ACN higher than 5% would result in a poor adsorption of the protein by T2.

2) Evaluation of the flow rate: Based on results obtained in 1), a solvent mixture containing 5% of ACN in 1.4 M ammonia was used to test three different flow rates: 100, 200, and 500 μL/min. Statistical results showed that about 90% of RNase and 100% of BDK were trapped even at the highest tested flow rate.

A complementary experiment was carried out using a real cleavage product. The least-retained cleavage product for all isoforms of RNase-A is the A-fragment 1-25 coming from the cleavage of IF2. This product was purified and

loaded into T2 under the experimental conditions that were described above. The recovery for this cleavage product was calculated at 95.2% (See Figure 77). From obtained results for the parameters evaluated for T2, we conclude that in order to obtain a good recovery in T2, it was necessary to use no more than 5% of ACN with a possible a flow rate up to 500 μL/min.

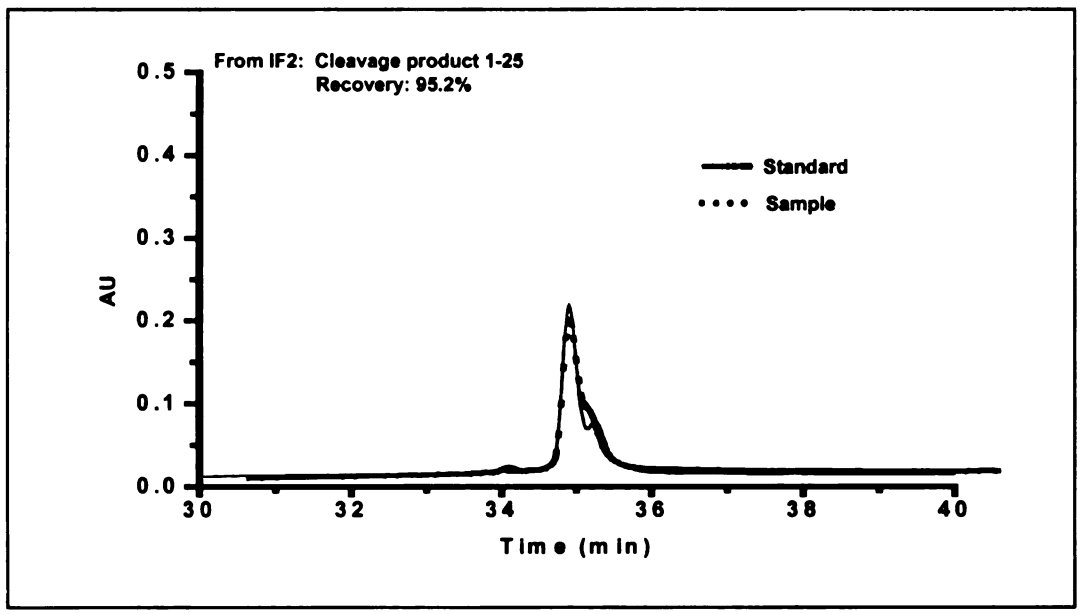


Figure 77. Chromatogram of cleavage product 1-25 from IF2 for RNase-A. The fact that the peaks for the 'standard' and the 'sample' cleavage product are almost overlapped indicates that recovery of the sample is close to 100% (in this example, 95.2%).

3.2c Trap 1 evaluation: Resin selection, load capacity efficiency, feasibility of partial reduction and cyanylation chemical reactions.

In this approach, trap 1 is never in contact with an extremely alkaline or acidic solution, which is why it is not essential that the resin in T1 be pH-resistant. Thus, several possible resins for partial reduction were tested under the same experimental conditions. RNase-A (aprox. 1 nmol) was loaded into trap 1, and a 0.02 M TCEP in 6 M guanidine/0.1 M citrate buffer pH = 3 solution

was injected to the trap. The solution was kept inside the column for 15 minutes and then trap 1 was connected immediately on-line to a narrowbore HPLC-column for a chromatographic separation of the partially-reduced species formed in T1. Chromatograms obtained with different resins were compared to those relating to the in-vial approach in order to select the most adequate resin for trap 1, as is shown in Figure 78.

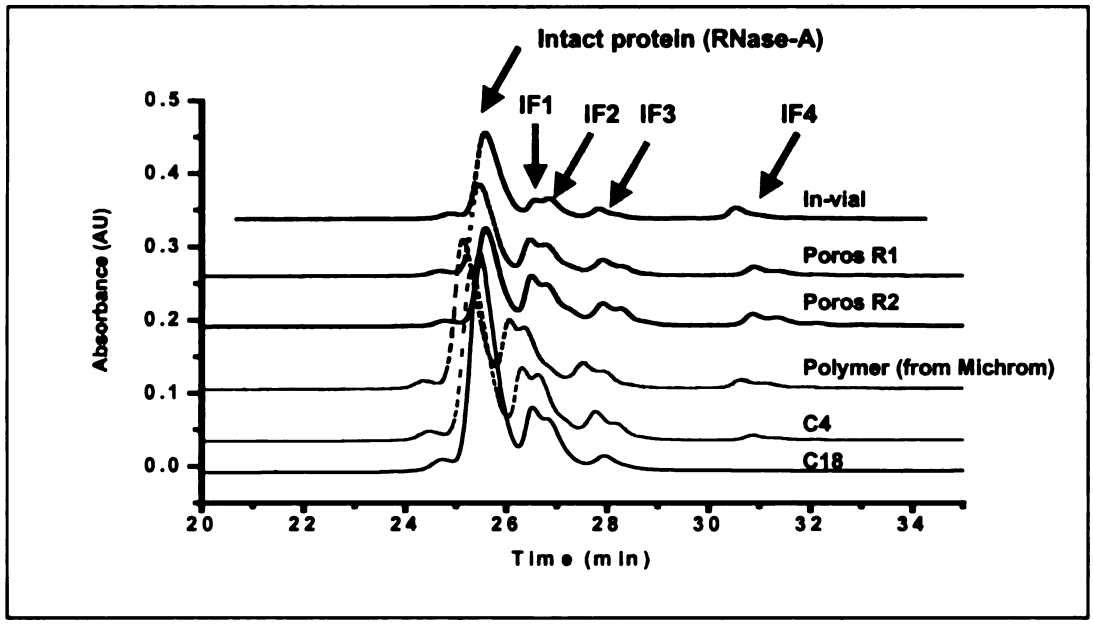


Figure 78. Comparison of chromatograms of a standard sample of partially reduced RNase-A as trapped by different stationary phase types. Intact RNase-A and its isoforms (IF1.. IF4) are labeled.

Because several different phases showed approximately the same efficiency for the partial reduction reaction, it was decided to use Poros R2 resin to fill the microbore column T1 in order to take advantage of its polymeric character and the low backpressure that this support offers, which would have the benefit of a lower overall pressure in the on-line system.

In order to evaluate the load capacity efficiency for T1 filled with Poros R2, a similar instrumental setup to the one shown in Figure 76 (with T1 instead of T2) was used. RNase-A (~ 1 nmol) and BDK (~ 5 nmol) were loaded under acidic conditions (pH = 3) at a flow rate of 50 μL/min (using the "load" pathway). In a second step, trap 1 was connected to the narrowbore column ("elution" pathway); retained proteins were eluted and chromatograms were obtained. The load capacity efficiency was evaluated from chromatograms by comparison with the ones of standard solutions of BDK and RNase-A. As expected from the loading capacity studies for trap 2, more than 95% of the proteins were retained in trap 1.

In order to optimize the partial reduction of a protein bound to the polymeric stationary phase, about 1 nmol of RNase-A was loaded and bound to trap 1, then TCEP was added to the trap at different concentrations. The reaction was allowed to proceed for specific reaction times; then, chromatograms of the reaction mixture were obtained at each partial reduction condition and the profiles were analyzed. Figure 79 shows representative chromatograms of this analysis in a comparison between the in-vial and the on-line approach. A concentration of about 20 nmol/μL for a reaction time of 15 minutes seems to be a good starting point for the partial reduction of a protein. It should be pointed out that using either the in-vial approach or the on-line approach, the experimental conditions for partial reduction could be modified according to the protein structure, which determines the accessibility of the reducing reagent to the disulfide bonds.

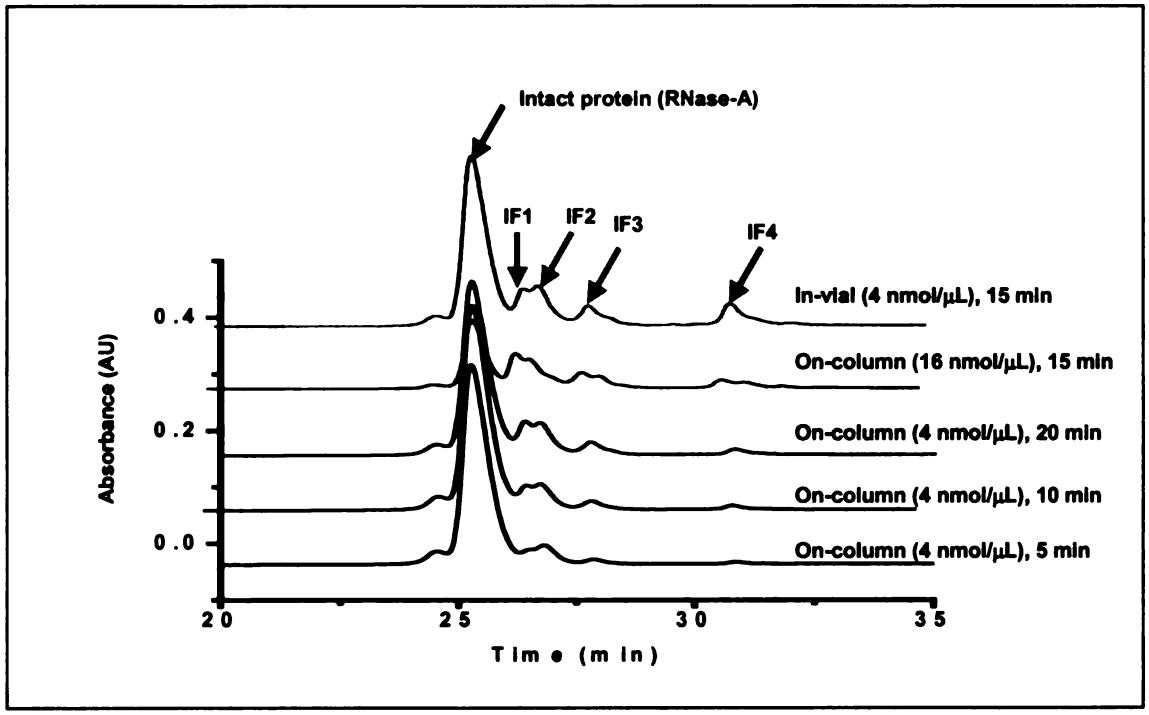


Figure 79. Comparison of HPLC chromatograms of partial reduction reaction mixture of 1 nmol of RNase-A on microbore column filled with Poros R2.

Cyanylation was evaluated in a quite simple manner. Each of the single reduced isoforms of the RNase-A were loaded, in independent experiments, into T1. CDAP solution was added to the trap for the cyanylation reaction to take place for 15 minutes. After that, T1 was connected on-line with the narrowbore-HPLC column and chromatograms were obtained. In Figure 80, three different chromatograms are shown. One corresponds to the single reduced IF4 (bottom panel) without cyanylation treatment, the other two chromatograms correspond to the same isoform under two different cyanylation treatments. From these results, it was determined that a solution of 0.2 M CDAP prepared in 6 M guanidine in 0.1 M citrate buffer allows a complete and quantitative cyanylation reaction.

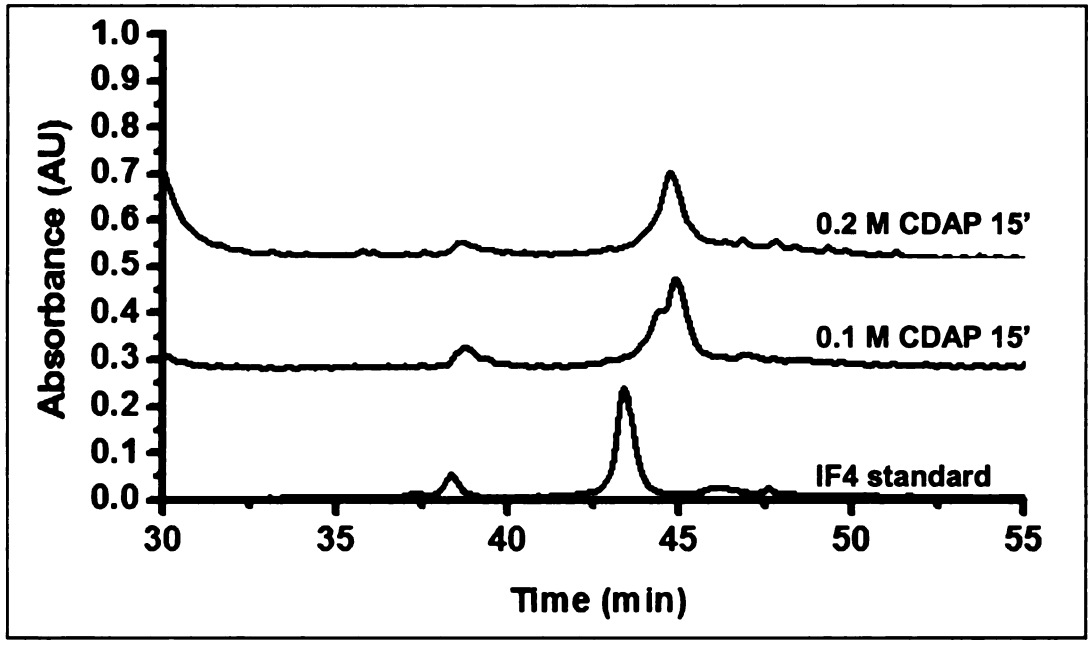


Figure 80. Chromatograms for cyanylation on-line with the two-column system. About 1 nmol of RNase-A was treated with CDAP at different concentrations during 15 minutes. From comparing the chromatograms, it is possible to observe that using a concentration of 0.2 M of CDAP drives the cyanylation to completion.

3.2d Trap 1 evaluation: Desorption efficiency

For successful PRCC methodology (as stated in 3.2a), after the partial reduction and cyanylation is carried out in trap 1, the products formed must be eluted quantitatively from T1. According to the instrumental setup shown in Figure 74 species eluted from T1 go directly to the reactor (R) where they are cleaved; the resulting cleavage products are captured in trap 2. For efficient cleavage, species desorbed from T1 must spend a certain time (residence time) in the reactor in order for the cleavage reaction to reach completion. For this experiment, the cleavage reagent solution (that is pumped into the mixer from a independent reagent bottle) was set to a continuous flow rate of 400 μL/min to allow the cleavage reaction to take place in the reactor in about 10 minutes of

residence time for species desorbed from T1. In order efficiently trap the cleavage products, the concentration of ACN in the eluent coming from the mixer to trap 2 has to be lower than 5%, otherwise cleavage products will flow out of T2. Thus, a solvent flow program was used that would desorb the proteins from trap 1 while providing a continuous constant flow of the cleavage reagent. The flow program for desorption from T1 was set as listed in Table 10, while different percentages of ACN were tested to evaluate the desorption efficiency of proteins from T1.

Table 10. Flow program to desorb isoforms from trap 1.

T (min)	*F (µL/min)*
0 – 0.5	50
0.5 – 1.0	40
1.0 – 30.0	20
30.0	0.0

For desorption efficiency experiments, RNase-A (~ 1 nmol) and BDK (~ 5 nmol) were loaded ontro trap 1 under acidic conditions (pH = 3). T1 was connected to the system as shown in Figure 74 (PRCC-system), and independent experiments using different percentages of ACN were carried out. After every experiment, T1 was removed from the PRCC-system and connected to a HPLC system (see Figure 81) to obtain the corresponding chromatogram. It was expected that after the desorption step, the total absence of chromatographic peaks for BDK and RNase-A would indicate that the desorption step had been successful.

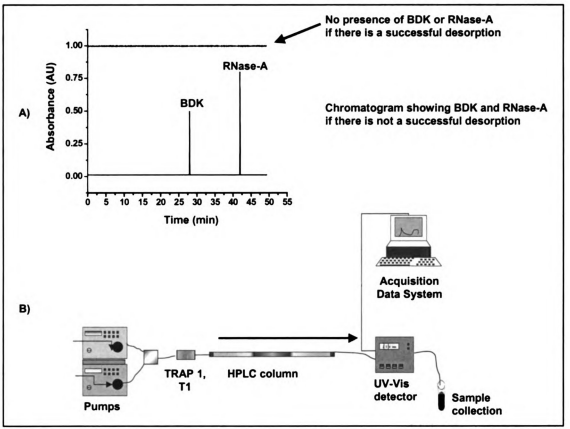


Figure 81. Conceptualized HPLC chromatogram (see A) showing that after desorption experiments, when T1 is connected to a HPLC system (see B) there is an absence of residual proteins in T1, which indicates a successful desorption step.

After several experiments (data not shown), it was determined that the use of 90% or 100% ACN (in water containing 0.1% trifluoroacetic acid) was not efficient for the desorption of the protein. However, the use of 80% ACN was optimum for a quantitative elution of the protein isoforms from trap 1.

3.2e Reactor Evaluation: Protein Transfer and CN-Protein Cleavage

After the partial reduction and cyanylation in trap 1, cyanylated isoforms are desorbed from it, cleaved in the reactor, and trapped by means of T2. Thus,

it was necessary to evaluate the efficiency of the transfer from T1 to T2. For this purpose, using a system setup as shown in Figure 82, RNase-A (~1 nmol) and BDK (~5 nmol) were loaded into trap 1. Using the flow program from Table 10 and a flow rate of 400 μL/min for the cleavage solution (1.4 M ammonia), proteins were desorbed from trap 1 and transferred through the reactor in a continuous flow to trap 2.

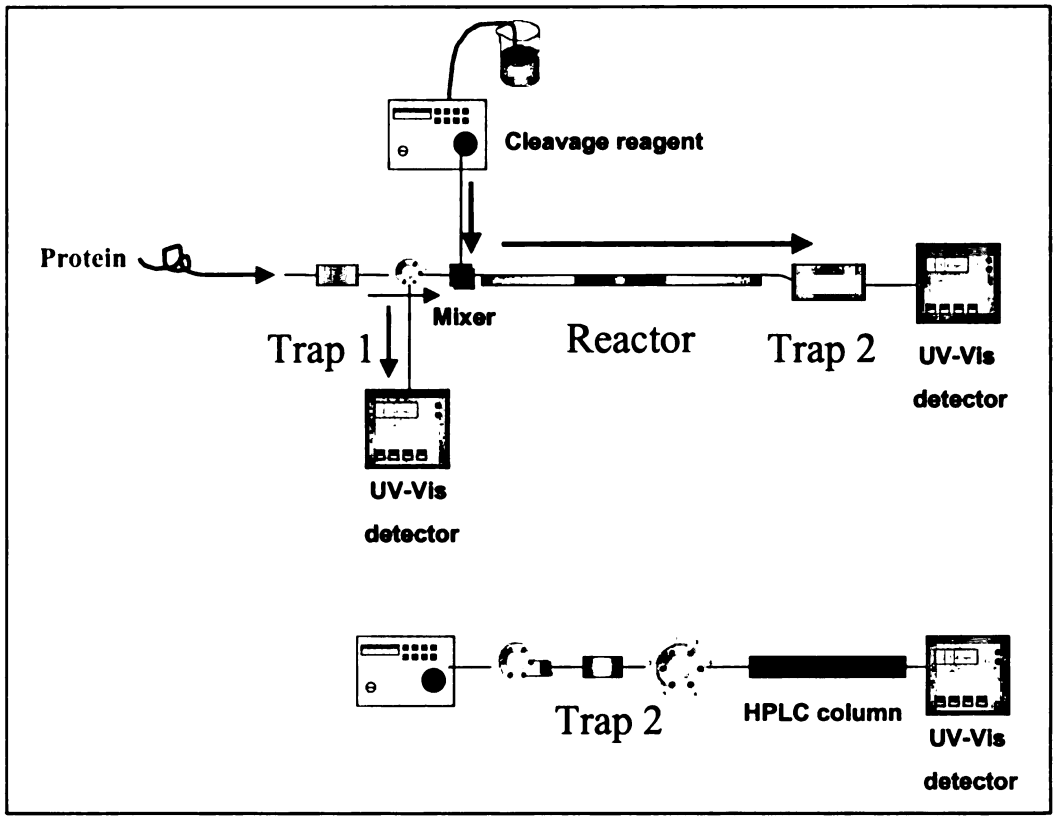


Figure 82. Top: Setup for the PRCC methodology using a two-column on-line system. Bottom: Schematic diagram indicating that trap 2 can be connected to a separation system for analysis of the cleavage products formed in the on-line system.

After this, trap 2 was connected on-line with a separation system as illustrated on the bottom scheme in Figure 82. Chromatograms were obtained

and recovery was calculated. In both cases, recovery was above 90% for RNase-A (92 %) and BDK (95 %).

For testing the on-line system for the cleavage reaction, preformed IF3 was loaded into trap 1 and, using the experimental steps described previously, IF3 was transferred from trap 1 to trap 2.

It was hypothesized that during this transfer step the cleavage reaction would take place cleaving IF3 at the CN-cysteines. Thus, after the transfer from T1 to T2, cleaved isoforms of IF3 would be the species trapped in T2. Trap 2 was washed with water and 0.1 M TCEP was added; the trap was immersed in a water bath at 30 °C for 30 min for the total reduction reaction. Then, trap 2, containing the released cleavage products, was connected to a separation system and a gradient solvent program was started to obtain the chromatogram shown in Figure 83. Chromatographic fractions were collected and cleavage products were identified by means of MALDI-MS. In Figure 83, a comparison of chromatograms for products of the in-vial and on-line approaches is shown. In general, both chromatograms show similar patterns; thus, the on-line approach demonstrated its feasibility for the PRCC methodology.

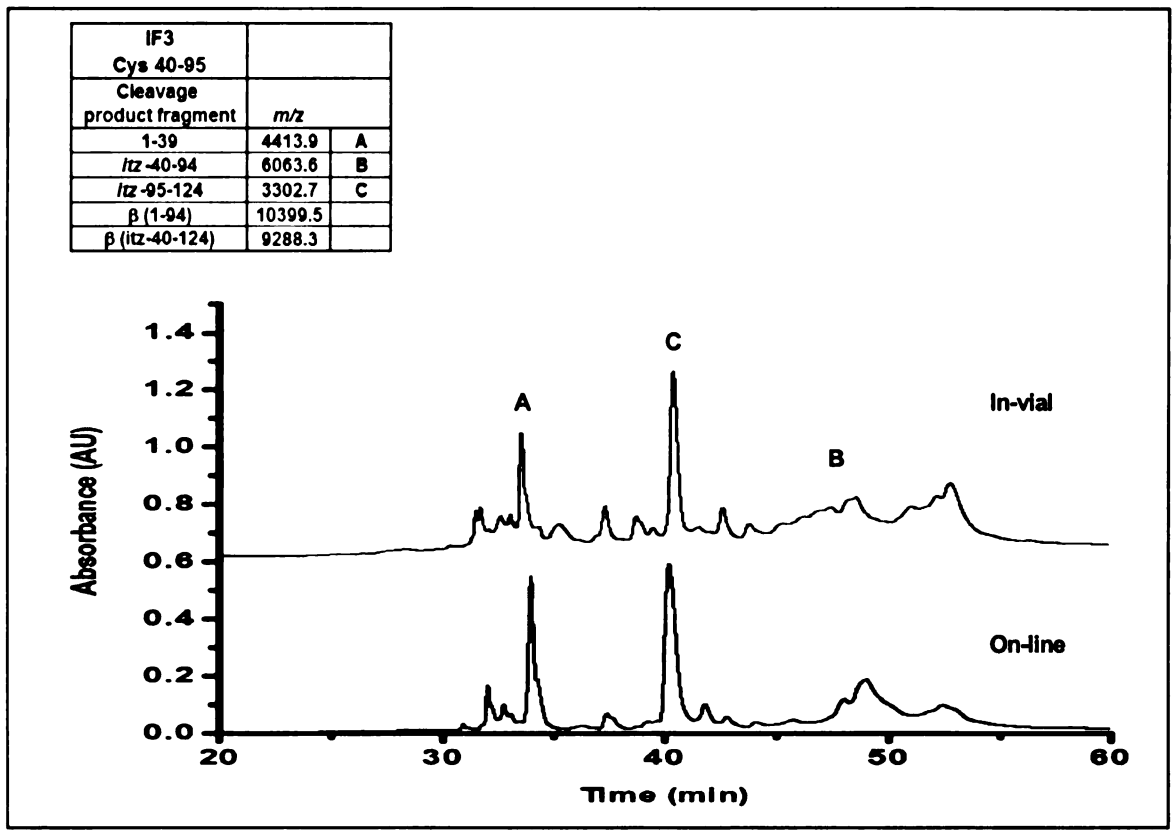


Figure 83. Comparison of chromatograms showing cleavage products for IF3 with in-vial and on-line approaches under similar experimental conditions (experimental conditions were similar to the described in section 3.2d). Table inset shows the *m/z* values for cleavage products for IF3.

3.2f Overall Evaluation of the on-line system: The on-line PRCC methodology

For the overall evaluation of the system, the next steps were followed:

1) Load sample load into trap 1

2) Add 0.02 M TCEP in 6M guanidine-0.1M citrate buffer at pH = 3 for partial reduction (15')

3) Wash trap to remove TCEP

4) Add 0.2 M CDAP in 6M guanidine-0.1M citrate buffer at pH = 3 for cyanylation (15')

5) Wash trap to remove CDAP

6) Cleavage in the reactor during transfer from Trap 1 to Trap 2 while pumping 1 M NH_4OH (or 1 M methylamine) through the mixer (40')

7) Wash Trap

8) Add 0.1M TCEP in 6M guanidine-0.1M citrate buffer at pH 3, for total reduction in Trap 2 (30 min at 37°C)

9) Connect Trap 2 to the HPLC-system and elute cleavage products using a solvent gradient program of ACN (0.1% TFA)/Water (0.1%TFA).

10) Collect fractions from the HPLC system and acquire MALDI-MS spectra.

RNase-A (~ 1 nmol) was loaded into the system (Trap 1) and all steps described above were carried out. MALDI-MS spectra were acquired for the partial separation of cleavage products eluted from Trap 2. Two independent experiments were carried out using ammonia or methylamine as cleavage reagents. At the top of Figure 84, a MALDI mass spectrum is shown from one of the fractions collected after the whole PRCC methodology was applied on-line using NH_3 as the nucleophile. At the bottom of the Figure 84 MALDI-MS signals are shown for cleavage products A and B for IF4 coming from experiments using ammonia or methylamine as cleavage reagents.

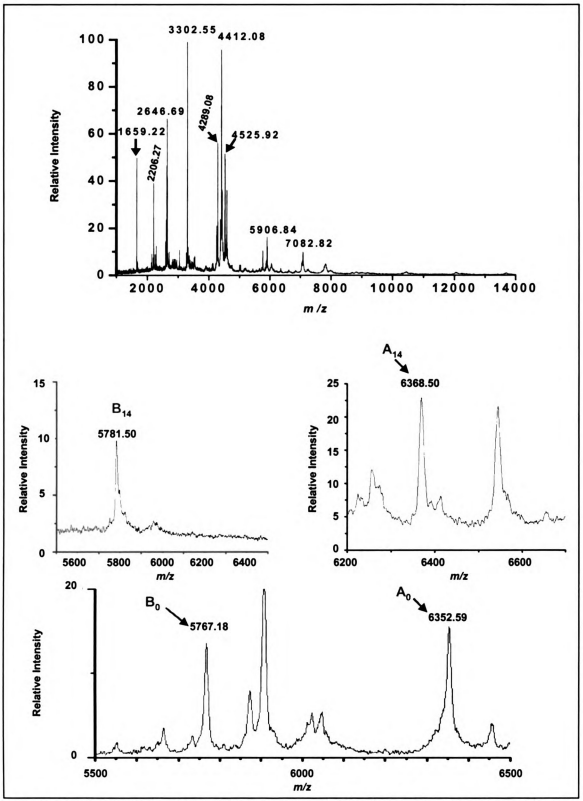


Figure 84. Cleavage products produced using an on-line system for PRCC with NH_3 as cleavage reagent. Top: Peaks representing cleavage products are identified with an arrow. Bottom: Cleavage products identified for IF4 with different nucleophiles, ammonia (Ao, Bo) or methylamine (A_{14}, B_{14}). Experiments with different nucleophiles were carried out independently.

From the MALDI mass spectra, cleavage products were identified as usual using the in-vial approach.

In Table 11 is presented a summary of the cleavage products that could be assigned from the MALDI spectra obtained from different collected fractions using ammonia as the cleavage reagent. It is apparent that not all expected cleavage products coming from singly reduced/cyanylated isoforms were identified in these experiments; however some other cleavage products coming from double and triply/cyanylated isoforms could be distinguished and assigned. Data obtained from these experiments with the 'on-line' approach were used as input for the two algorithms previously developed in this laboratory by former group members [41, 42]. Both algorithms turned out the correct linkage for the model protein. Thus, the 'on-line' approach can be considered as a viable alternative for the application of the PRCC methodology in characterizing cystinyl proteins.

Table 11. Summary of the cleavage products assigned for an 'on-line' experiment using ammonia. Fragments indicated with * were produced by double or triply reduced and cyanylated isoforms of RNase-A.

IFs	Detected	Fragment	Calculated *m/z* values (Da)
IF1	✓ Y	1-64	7083.90
IF1	✓ Y	72-124	5906.50
IF1	N ✗	65-71	789.80
IF2	✓ Y	1-25	2706.87
IF2	✓ Y	84-124	4526.00
IF2	N ✗	26-83	6547.30
IF3	✓ Y	1-39	4413.90
IF3	✓ Y	40-94	6063.00
IF3	✓ Y	95-124	3302.70
IF4	✓ Y	110-124	1669.80
IF4	✓ Y	1-57	6353.10
IF4	✓ Y	58-109	5767.00
*	✓ Y	26-40	1686.96
*	✓ Y	40-58	1751.05
*	✓ Y	58-84	2901.22
*	✓ Y	72-110	4290.80
*	✓ Y	72-94	2847.00

4.- CONCLUSIONS

During use of the 'on-line' approach, there are no sample transfers or drying steps involved; thus, sample manipulation is minimal and sample losses are minimized. Because sample losses are minimized, a lower amount of protein is necessary as a starting material.

The 'on-line' approach takes only half the time required for the regular in-vial approach (see Table 12).

Table 12. Comparison of sample processing time required for in-vial and on-line approaches.

Time evaluation		
- Time associated with the complete determination of the location of disulfide bonds:		
	In-Vial	**On-line**
1A) Load Sample	**1'**	**10'**
1) Partial reduction	**15'**	**15'**
2) Cyanylation	**15'**	**15'**
3) Chromatography	**30-90'**	**-**
4) Solvent removal	**120'**	**-**
5) Cleavage	**60'**	**10'**
6) Solvent removal	**20'**	**-**
7) Total reduction	**30'**	**30'**
8) Optional: Chromatography	**30-90'**	**60'**
Estimated total time:	**6 - 8 hrs**	**3 - 4 hrs**

Trap 2, containing the cleavage products, can be adapted easily to a separation system that could be coupled to a MS system, such a LC-MS.

5.- REFERENCES

[1] N. F. Visser, H. Lingeman, H. Irth, On-line SPE-RP-LC for the determination of insulin derivatives in biological matrices, *J. Pharm. Biomed. Anal.* 32 (2003) 295-309.

[2] L. E. Vera-Avila, E. A. Cazares-Ibanez, R. Covarrubias-Herrera, E. Camacho-Frias, On-line methodology for ultratrace determination of polynuclear aromatic hydrocarbons in water, *Rev. Int. Cont. Amb.* 18 (2002) 5-16.

[3] L. E. Vera-Avila, J. L. Meraz-Lira, P. Padilla-Cortes, On line methodology for the determination of trace levels of the herbicides 2,4-D and 2,4-DB in water, *Revista Internacional de Contaminacion Ambiental* 13 (1997) 63-71.

[4] S. Salado C, L. E. Vera-Avila, Online solid-phase extraction and high-performance liquid chromatographic determination of chlorthalidone in urine, *J. Chromatogr., B* 690 (1997) 195-202.

[5] E. Papadopoulou-Mourkidou, J. Patsias, E. Papadakis, A. Koukourikou, Use of an automated on-line SPE-HPLC method to monitor caffeine and selected aniline and phenol compounds in aquatic systems of Macedonia-Thrace, Greece, *Fresenius J. Anal. Chem.* 371 (2001) 491-6.

[6] P. Campins-Falco, R. Herraez-Hernandez, A. Sevillano-Cabeza, Column-switching techniques for high-performance liquid chromatography of drugs in biological samples, *J. Chromatogr.* 619 (1993) 177-190.

[7] P. Campins-Falco, R. Herraez-Hernandez, A. Sevillano-Cabeza, Column-switching techniques for screening of diuretics and probenecid in urine samples, *Anal. Chem.* 66 (1994) 244-248.

[8] L. E. Vera-Avila, L. Rangel-Ordonez, R. Covarrubias-Herrera, Evaluation and characterization of a commercial immunosorbent cartridge for the solid-phase extraction of phenylureas from aqueous matrices, *J. Chromatogr. Sci.* 41 (2003) 480-488.

[9] N. Takahashi, Y. Takahashi, N. Ishioka, B. S. Blumberg, F. W. Putnam, Application of an automated tandem high-performance liquid chromatographic system to peptide mapping of genetic variants of human serum albumin, *J. Chromatogr.* 359 (1986) 181-191.

[10] N. Takahashi, N. Ishioka, Y. Takahashi, F. W. Putnam, Automated tandem high-performance liquid chromatographic system for separation of extremely complex peptide mixtures, *J. Chromatogr.* 326 (1985) 407-418.

[11] M. M. Bushey, J. W. Jorgenson, Automated instrumentation for comprehensive two-dimensional high-performance liquid chromatography of proteins, *Anal. Chem.* 62 (1990) 161-167.

[12] G. J. Opiteck, K. C. Lewis, J. W. Jorgenson, R. J. Anderegg, Comprehensive on-line LC/LC/MS of proteins, *Anal. Chem.* 69 (1997) 1518-1524.

[13] G. J. Opiteck, J. W. Jorgenson, Two-dimensional SEC/RPLC coupled to mass spectrometry for the analysis of peptides, *Anal. Chem.* 69 (1997) 2283-2291.

[14] K. Wagner, T. Miliotis, G. Marko-Varga, R. Bischoff, K. K. Unger, An automated on-line multidimensional HPLC system for protein and peptide mapping with integrated sample preparation, *Anal. Chem.* 74 (2002) 809-820.

[15] R. E. Majors, Multidimensional and comprehensive liquid chromatography, *LC-GC North America* 23 (2005) 1074-1076, 1078, 1080-1082, 1084.

[16] J. C. Giddings, Concepts and comparisons in multidimensional separation, *J. High Resolut. Chromatogr.* 10 (1987) 319-323.

[17] Y. Mao, P. W. Carr, Adjusting selectivity in liquid chromatography by use of the thermally tuned tandem column concept, *Anal. Chem.* 72 (2000) 110-118.

[18] Y. Mao, P. W. Carr, Application of the thermally tuned tandem column concept to the separation of several families of environmental toxicants, *Anal. Chem.* 72 (2000) 2788-2796.

[19] Y. Mao, P. W. Carr, Separation of selected basic pharmaceuticals by reversed-phase and ion-exchange chromatography using thermally tuned tandem columns, *Anal. Chem.* 73 (2001) 4478-4485.

[20] E. Edwards, J. thomas-Oates, Hyphenating liquid phase separation techniques with mass spectrometry: on-line or off-line, *Analyst* 130 (2005) 13-17.

[21] B. K. Choi, Y. M. Cho, S. H. Bae, C. C. Zoubaulis, Y. K. Paik, Single-step perfusion chromatography with a throughput potential for enhanced peptide detection by matrix-assisted laser desorption/ ionization-mass spectrometry, *Proteomics* 3 (2003) 1955-1961.

[22] Y. Shen, N. Tolic, C. Masselon, L. Pasa-Tolic, D. G. Camp, 2nd, K. K. Hixson, R. Zhao, G. A. Anderson, R. D. Smith, Ultrasensitive proteomics using high-efficiency on-line micro-SPE-nanoLC-nanoESI MS and MS/MS, *Anal. Chem.* 76 (2004) 144-154.

[23] R. Xiang, C. Horvath, J. A. Wilkins, Elution-modified displacement chromatography coupled with electrospray ionization-MS: on-line detection of trace peptides at low-femtomole level in peptide digests, *Anal. Chem.* 75 (2003) 1819-1827.

[24] Y. Jiang, C. S. Lee, On-line coupling of micro-enzyme reactor with micro-membrane chromatography for protein digestion, peptide separation, and protein identification using electrospray ionization mass spectrometry, *J. Chromatogr. A* 924 (2001) 315-322.

[25] L. Wang, D. L. Smith, Downsizing improves sensitivity 100-fold for hydrogen exchange-mass spectrometry, *Anal. Biochem.* 314 (2003) 46-53.

[26] S. Hara, V. Katta, H. S. Lu, Peptide map procedure using immobilized protease cartridges in tandem for disulfide linkage identification of neu differentiation factor epidermal growth factor domain, *J. Chromatogr. A* 867 (2000) 151-160.

[27] G. Massolini, E. Calleri, Immobilized trypsin systems coupled on-line to separation methods: recent developments and analytical applications, *J. Sep. Sci.* 28 (2005) 7-21.

[28] G. Marie, L. Serani, O. Laprevote, An on-line protein digestion device for rapid peptide mapping by electrospray mass spectrometry, *Anal. Chem.* 72 (2000) 5423-5430.

[29] J. Wu, J. T. Watson, A novel methodology for assignment of disulfide bond pairings in proteins, *Protein Sci.* 6 (1997) 391-398.

[30] ASTM D 1193-99e1, "Standard Specification for Reagent Water", ASTM International.

[31] M. Cadene, B. T. Chait, A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins, *Anal. Chem.* 72 (2000) 5655-5658.

[32] M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Matrix-assisted ultraviolet laser desorption of non-volatile compounds, *Int. J. Mass Spectrom. Ion Processes* 78 (1987) 53-68.

[33] F. Hillenkamp, U. Bahr, M. Karas, B. Spengler, Mechanisms of laser ion formation for mass spectrometric analysis, *Scanning Microscopy, Supplement* 1 (1987) 33-39.

[34] R. L. Cunico, K. M. Gooding, T. Wehr, Basic HPLC and CE of biomolecules, Bay Bioanalytical Laboratory, Richmond, CA, 1998.

[35] L. R. Snyder, J. J. Kirkland, Introduction to modern liquid chromatography, Wiley, New York, NY., 1974.

[36] E. Katz, Handbook of HPLC, Marcel Dekker, New York, 1998.

[37] L. E. Vera-Avila, A. M. Nunez-Gaytan, M. R. Covarrubias-Herrera, Trace level chromatographic determination of priority pollutant chloro- and nitrophenols in water with online sample preparation, *Int. J. Environ. Anal. Chem.* 71 (1998) 147-164.

[38] Hamilton. PRP-3, 300 Å Reversed Phase HPLC Columns, (accessed 2003, 2005).

[39] J. Kyte, R. F. Doolittle, A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.* 157 (1982) 105-132.

[40] N. B. Afeyan, N. F. Gordon, I. Mazsaroff, L. Varady, S. P. Fulton, Y. B. Yang, F. E. Regnier, Flow-through particles for the high-performance liquid chromatographic separation of biomolecules: perfusion chromatography, *J. Chromatogr.* 519 (1990) 1-29.

[41] W. Wu, W. Huang, J. Qi, Y. T. Chou, E. Torng, J. T. Watson, 'Signature sets', minimal fragment sets for identifying protein disulfide structures with cyanylation-based mass mapping methodology, *J. Proteome Res.* 3 (2004) 770-777.

[42] J. Qi, W. Wu, C. R. Borges, D. Hang, M. Rupp, E. Torng, J. T. Watson, Automated data interpretation based on the concept of "negative signature mass" for mass-mapping disulfide structures of cystinyl proteins, *J. Am. Soc. Mass Spectrom.* 14 (2003) 1032-1038.

CHAPTER 5. CONCLUSIONS AND FUTURE WORK

1.- CONCLUSIONS

During the implementation of an 'on-line' system for the Partial Reduction/Cyanylation CN-induced cleavage-based mass mapping methodology for cystinyl proteins, it was realized that the cleavage reaction needed to be optimized. Following a quantitative assessment of the cleavage reaction using ammonia, primary amines were proposed as alternative nucleophiles. Methylamine was found to be superior to ammonia as the cleavage nucleophile in promoting the CN-induced cleavage reaction during sample preparation for disulfide mass mapping. However, increasing the concentration of ammonia to 5 M (beyond the previously used 1 M) in combination with low temperature (2° C instead of room temperature) and short reaction time increased the cleavage product yields substantially [1]. Thus, it seems that low temperature (2 °C) and a short reaction time (10 min) minimize alkali-induced damage of the cleavage products and other side reactions that occur during longer incubation times in the original protocol (1 hour or more). Good repeatability (less than 10% relative standard deviation) was observed in the yield of cleavage products from ~100-1000 pmol of starting material.

After the optimization of the experimental conditions for the cleavage reactions, the use of two or more homologous nucleophiles were recognized as

an additional tool to facilitate the identification of cleavage products from mass spectra. Mixtures of methylamine/ammonia as nucleophiles produce homologous species of certain cleavage products that a are represented by pairs of mass spectral peaks separated by 14 mass units, which helps the analyst distinguish these structurally diagnostic species, especially in cases of low-intensity signals and/or in the presence of potentially interfering chemical noise.

The results of evaluating an 'on-line' system for the chemical modification of cystinyl proteins show great promise for the implementation of an automatic system in the future where small amounts of starting material (few pmols) can be analyzed without transfer steps in relatively short experimental time.

2.- FUTURE WORK

As it was stated in Chapter 1, disulfide bonds are one of the key aspects related to folding among cystinyl proteins and, consequently, their determination is important for the understanding of the chemical structure and function of these kinds of proteins, which are emerging as an important class of bioactive proteins [2, 3]. Currently, proteomics is focused on the identification of proteins based on a database search involving mass spectrometry; only minor attention is paid to the determination of the disulfide proteome. However, it is expected that proteomic research will start to focus on the state of disulfide linkages in cystinyl proteins; that is why the cyanylated-based mass mapping methodology may gain

more importance in a near future; in fact, one feature that currently distinguishes this methodology is its successful application to cystinyl proteins containing adjacent cysteines [4-6].

Thus, the future of cyanylation-based mass mapping methodology should rely on the development of a completely automated system, where the starting point is the loading of the cystinyl protein and the end point is the report stating the disulfide linkage. In order to reach this goal, I propose the following projects:

1) Evaluation of the partial reduction experimental conditions: In this thesis, the experimental conditions for the cleavage reaction were optimized. However, the first step in the methodology is the partial reduction and cyanylation of the cystinyl protein. It seems that cyanylation does not represent a major problem, however, sometimes partial reduction of a protein can be difficult to accomplish due to the protein structure being so knotted that access of the reducing agent is limited. The use of a step-wise partial reduction/alkylation (or cyanylation)-sequence [7-11] in which the protein would be reduced and modified (alkylated or cyanylated) sequentially should be evaluated for possible incorporation into an 'on-line' system. The protein should be alkylated after the first partial reduction step, and cyanylated after a second reduction step; in this way, two disulfide linkages could be determined. However, it would be necessary to determine whether alkylated cysteines with certain chemical groups are affected during the cleavage step.

Another choice for dealing with the partial reduction step could be the use of electrochemistry; currently, there is not too much scientific information on this topic, and it has to be explored.

2) Development of new software: After the development of the CN-based mass mapping methodology [12], it was evident that experimental data had to be analyzed more efficiently, thus, the Negative Signature Algorithm (NSA) [13, 14] was developed to help in the data analysis for the determination of the disulfide linkage; however, because it was the first version, this program lacked flexibility because it considered that the methodology was carried out under almost ideal conditions. Sometimes results from this algorithm can be misleading if they are not analyzed cautiously, especially from mass spectra of low signal strength. Fragment Sets Algorithm (FSA) [15] was the second software developed; the nice feature of this software is that it can deal with imperfect experimental data providing an output list with a ranking of the most probable disulfide structures for the protein sample. For both algorithms, reliable results depend on identification of the cleavage fragments in the mass spectra.

In this thesis, more dependable identification of cleavage products is proposed by the use of two homologous nucleophiles for labeling of the cleavage products. Thus, a new software needs to be developed where the input is raw experimental data from a sample containing labeled cleavage products. This algorithm must be able to identify cleavage products by means of the mass

difference between pairs of mass spectral peaks representing the labeled (homologous) cleavage products. After this, the algorithm could be linked to the NSA or to FSA for the determination of the disulfide linkage.

3) Integration of the CN-based methodology into a totally automated system: As it was sketched above, one of the main goals would be the development of an automatic system in which the cystinyl protein sample could be loaded, chemically derivatized, and cleaved. Cleavage products analyzed by mass spectrometry for automatic data interpretation by the use of a new algorithm could give the correct disulfide linkage. The on-line system proposed in this thesis can be the starting point to reach this goal.

3.- REFERENCES

[1] J. L. Gallegos-Perez, L. Rangel-Ordonez, S. R. Bowman, C. O. Ngowe, J. T. Watson, Study of primary amines for nucleophilic cleavage of cyanylated cystinyl proteins in disulfide mass mapping methodology, *Anal. Biochem.* 346 (2005) 311-319.

[2] H. Yano, S. Kuroda, B. B. Buchanan, Disulfide proteome in the analysis of protein function and structure, *Proteomics* 2 (2002) 1090-1096.

[3] J. J. Gorman, T. P. Wallis, J. J. Pitt, Protein disulfide bond determination by mass spectrometry, *Mass Spectrom. Rev.* 21 (2002) 183-216.

[4] J. Wu, J. T. Watson, Optimization of the cleavage reaction for cyanylated cysteinyl proteins for efficient and simplified mass mapping, *Anal. Biochem.* 258 (1998) 268-276.

[5] T. P. Wallis, J. J. Pitt, J. J. Gorman, Identification of disulfide-linked peptides by isotope profiles produced by peptic digestion of proteins in 50% (18)O water, *Protein Sci.* 10 (2001) 2251-2271.

[6] T. P. Wallis, C. Y. Huang, S. B. Nimkar, P. R. Young, J. J. Gorman, Determination of the disulfide bond arrangement of dengue virus NS1 protein, *J. Biol. Chem.* 279 (2004) 20729-20741.

[7] S. Sechi, B. T. Chait, Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification, *Anal. Chem.* 70 (1998) 5150-5158.

[8] V. Schnaible, S. Wefing, A. Bucker, S. Wolf-Kummeth, D. Hoffmann, Partial reduction and two-step modification of proteins for identification of disulfide bonds, *Anal. Chem.* 74 (2002) 2386-2393.

[9] J. E. Hale, J. P. Butler, V. Gelfanova, J. S. You, M. D. Knierman, A simplified procedure for the reduction and alkylation of cysteine residues in proteins prior to proteolytic digestion and mass spectral analysis, *Anal. Biochem.* 333 (2004) 174-181.

[10] U. Goransson, D. J. Craik, Disulfide mapping of the cyclotide kalata B1. Chemical proof of the cystic cystine knot motif, *J. Biol. Chem.* 278 (2003) 48188-48196.

[11] U. Goransson, A. M. Broussalis, P. Claeson, Expression of Viola cyclotides by liquid chromatography-mass spectrometry and tandem mass spectrometry sequencing of intercysteine loops after introduction of charges and cleavage sites by aminoethylation, *Anal. Biochem.* 318 (2003) 107-117.

[12] J. Wu, J. T. Watson, A novel methodology for assignment of disulfide bond pairings in proteins, *Protein Sci.* 6 (1997) 391-398.

[13] J. Qi, W. Wu, C. R. Borges, D. Hang, M. Rupp, E. Torng, J. T. Watson, Automated data interpretation based on the concept of "negative signature mass" for mass-mapping disulfide structures of cystinyl proteins, *J. Am. Soc. Mass Spectrom.* 14 (2003) 1032-1038.

[14] J. Qi, J. Wu, G. A. Somkuti, J. T. Watson, Determination of the disulfide structure of sillucin, a highly knotted, cysteine-rich peptide, by cyanylation/cleavage mass mapping, *Biochemistry* 40 (2001) 4531-4538.

[15] W. Wu. *Applications and computational considerations of the cyanylation-based disulfide mapping methodology for cystinyl proteins*, PhD Thesis, Michigan State University, East Lansing, MI, December 2003.