Development of separatory and mass spectrometric methods for the characterization of marine metallothioneins

Thompson, J. A. J.; Boyd, R. K.; Pleasance, S.; Thibault, P.

Publisher's version / Version de l'éditeur:
https://doi.org/10.4224/23001236
IMB Technical Report, 1993

DEVELOPMENT OF SEPARATORY AND MASS SPECTROMETRIC METHODS FOR THE CHARACTERIZATION OF MARINE METALLOTHIONEINS

Issued as IMB Technical Report No. 70
NRCC No. 34893

December, 1993

Report prepared for:

Dr. J.A.J. Thompson,
Institute of Oceanographic Sciences,
Fisheries and Oceans Canada,
Sidney, B.C. V8L 4B2.

Report written by : **R.K. Boyd,**

based on experimental data and information supplied by :

S. Pleasance* and P. Thibault,
Analytical Chemistry Group,
Institute for Marine Biosciences,
National Research Council,
1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1.

*** Under contract from SCIEX, 55 Glencameron Road, Thornhill, Ontario L3T 1P2.**

Contents

A INTRODUCTION

Metallothioneins are small proteins of approximately 60 amino acids having a highly unusual composition, *viz.* no aromatic or heterocyclic amino acids while one third of the residues are cysteines. Another feature of these proteins, related to their cysteine content, is their high affinity for metal ions such as zinc or cadmium. Up to 7 moles of metal can be bound per mole of protein. These proteins have attracted considerable interest [1-30] in view of their *in vivo* inducibility in response to heavy metals and to a great variety of metabolites such as glucocorticoids, glucagon, interleukin 1, catecholamines, progesterone, estrogen, and ethanol, thus making them a convenient target for studying the regulation of gene expression [2]. Over the last forty years or so there has been a considerable amount of research dedicated to the further understanding of the role and functions of metallothioneins in various biological systems. It would be impossible to describe here the large body of literature in this field. Rather, the reader is referred to Volumes 158 and 205 of Methods in Enzymology (J.F. Riordan and B.L. Vallee Eds, Acad. Press, NY), which have conveniently assembled and assessed publications dealing with various aspects of metallothionein research. However, for purposes of convenience and to facilitate evaluation of the prospect of novel separatory and mass spectrometric methods for structural characterization of metallothioneins, it is convenient to briefly summarize here some of their functional and structural characteristics.

Detoxification and regulation of metal homeostasis

Metallothioneins are ubiquitous throughout the animal kingdom [3]. Since the first report on the isolation of cadmium-binding protein in equine kidney cortex some forty years ago [4], numerous investigations have been conducted on the extraction of metallothioneins from lower vertebrates, invertebrates, microorganisms [3,5,6] and even plants [7]. Although these complex biomolecules appear to be distributed widely throughout a variety of species, their specific biological role still triggers considerable debate. However, it is clear that their selectivity and specificity towards metals such as cadmium, zinc and copper implies that they have a significant role in metal metabolism and detoxification, and regulation of metal homeostasis.

Metallothioneins remain the only proteins known to contain cadmium in their native states, and as such it is not surprising that they were first proposed to have an important role in heavy metal detoxification processes [8]. A reduction in the toxicity of metal ions following the induction of metallothioneins in animals has long been recognized. Although the sequestration of intracellular cadmium by metallothionein might appear to allow the metal to be "stored" in a non-toxic form, the cadmium-bonded proteins themselves appear to be potent nephrotoxins targeting the proximal tubule. Intravenous injections of the cadmium-metallothionein complexes into rats [9,10], and additions to cultures of epithelial cells [11], have demonstrated their higher toxicity compared to that of equivalent or even higher doses of cadmium ion [12].

Thus intracellular metallothioneins appear to have a protective function against exposure to cadmium, whereas their extracellar counterparts are significantly more harmful.

In the light of the foregoing discussion it is difficult to believe that metallothionein can play solely a detoxification role. The most conspicuous biological feature of the metallothioneins is their inducibility by numerous agents including metal ions, hormones, growth factors, cytokines, vitamins, and antibiotics, as well as physical and chemical stress conditions. A list of these known inducing factors has been recently assembled by Kägi [13]. The large number of factors stimulating the biosynthesis of metallothionein is precisely the origin of the difficulty in attempting to define their biological functions. Evidently their primary structure homology (which has been conserved over evolutionary timescales), their ubiquitous occurrence, and their synthesis in development, regeneration and reproduction, are strong arguments suggesting that metallothioneins serve several specific metal-related cellular roles [14,15].

Additional insight into the biological functions of metallothioneins may be gained by considering their analogies with zinc enzymes and DNA-binding proteins. Their structural homologies with DNA-binding proteins in terms of conserved cysteine and histidine are remarkable [16], and have led to the proposal of a popular model referred to as a "zinc finger". Furthermore the abilities of zinc to adopt a particularly flexible

coordination sphere, and to assume multiple coordination geometries, are probably critical features pertaining to its biological roles. Metallothioneins can provide a reservoir for supplying zinc and copper in the biosynthesis of metalloenzymes and metalloproteins [17,18]. On the other hand the cell can exploit the coordination flexibility of zinc by incorporating it into structural sites of DNA-binding proteins where, by controlling the flow of zinc, the metallothioneins could serve a regulatory role in zinc-dependent processes in replication, transcription, and translation [19,20].

Classification of metallothioneins

The committee on the Nomenclature of Metallothioneins appointed at the General Discussion Session of the Second International Meeting on Metallothioneins and Other Low-molecular-Weight Binding Proteins in 1985, proposed the following three classes based on their structural characteristics [21,22]:

Class I: Polypeptides with locations of cysteine closely related to those in equine renal metallothionein.

Class II: Polypeptides with locations of cysteines only distantly related to those in equine renal metallothionein, such as yeast metallothioneins [23].

Class III: Atypical nontranslationally synthesized metal - thiolate polypeptides, such as cadystin [24], phytometallothionein, phytochelatin, and homophytochelatin [25-28].

Chemical Characterization of Metallothioneins.

This section summarizes a very large body of work, and is itself a summary of the account given by Kägi and Kojima [6], who give the original references.

Mammalian metallothioneins (MTs) have long been known to contain two isoforms, conveniently labelled MT-1 and MT-2, distinguishable by electrophoretic experiments. MT isoform preparations judged to be pure by the usual protein chemistry criteria are often heterogeneous with respect to the relative contents of Zn, Cd, Cu and other minor metallic constituents. Even here, however, the distribution of metal ions amongst the Cys binding sites is not completely random, and there is some bias in the distribution of Zn and Cd amongst the sites. (This observation is enlarged upon below). The total metal content seems to be essentially constant, however. Mammalian MTs contain seven moles of firmly bound Zn and/or Cd, while crab MTs contain six. There is some evidence that under some circumstances, particularly for Cu^{+}, higher molar ratios of metal: protein are observed. There is no evidence for lower (metal: MT) ratios *in vivo*.

In addition to their high cysteine content, most MTs also contain a large proportion of basic amino acids, usually Lys but occasionally Arg. Mammalian MTs contain no His, but this residue is found in MTs of some other species. Hydrophobic residues are conspicuous by their absence, or at least their rarity, in MTs. Thus, there are no aromatic residues in Class I MTs, though Phe, Tyr and His are found in some Class II MTs. Class III MTs are peculiar compounds, whose amino acid composition is restricted to Cys, Glu and Gly (occasionally replaced by ß-Ala), and are not discussed further in this report.

Most mammalian MTs contain 61 amino acid residues, but the two isoMTs from crab (*Scylla serrata*) contain 57 and 58 residues, respectively. The common feature of all Class I MTs is the preservation of Cys, *e.g.* in mammals all 20 Cys are invariant. Other MTs, such as that from crab which is of interest here, show a close correspondence of Cys positions with the mammalian MTs as illustrated in Table 1. In all cases a conspicuous feature is the appearance of Cys-X-Cys subsequences, where X is a residue other than Cys. Mammalian MTs also contain three Cys-Cys sub-sequences, though there are only two such sub-sequences in the crab isoMTs. In all mammalian MTs examined thus far, 35 of the 61 residues are invariant, amongst them the entire N-terminal segment comprising residues 1-7 and a ten-residue segment in the centre of the chain (residues 29-38). All vertebrate MTs whose primary sequences have been determined at the protein level have a blocked amino terminus, usually N-acetyl-Met. Both of the crab isoMTs have Pro as the N-terminal residue,

and this is not blocked (Table 1).

The primary sequences of MTs validate the early division of the mammalian MTs into the major sub-forms MT-1 and MT-2 (originally based upon differing net charges as determined by electrophoretic experiments). Thus, all mammalian MT-2 isoforms have an acidic residue in position 10 or 11 instead of the neutral residue found in the MT-1 isoform. This distinction is also true of the crab MTs (Table 1), in which the Val^{10} residue in MT-1 is replaced by ^{10}Asp in MT-2. However, primary sequence determinations have shown that the situation is more complicated than the original simple division into just two isoforms. Amino acid substitutions yield minor variants of either or both of the major charge-separable forms, *e.g.* rabbit liver MT-2 has been shown to contain three minor variants involving substitutions in six different positions.

A substantial effort has been devoted to the characterization of metal-binding sites in MTs. The major tools used in this effort have been ^{113}Cd NMR and X-ray crystallography, although optical spectroscopy has shown that the metal-sulphur bonding is of the metal-thiolate type, *i.e.* all Cys residues are deprotonated and serve as ligands to the bivalent metal ions with an average ratio of almost 3 ligands per metal ion (or exactly 3 in the case of crab MTs). The removal of metals from the proteins is achieved by exposure to strong chelating agents such as EDTA, or by acidification. Complete loss of Zn^{2+} and of Cd^{2+} from mammalian MTs occurs at pH3 and pH2, respectively. Demetallized MTs ("apoMTs") can be reconstituted by

exposure to the appropriate quantities of metal salts. The affinity of metal ions for the apoMTs increases in the order:

Zn(II) < Pb(II) < Cd(II) < [Cu(I), Ag(I), Hg(II), Bi(III)].

The conspicuous juxtaposition of the basic residues Lys and Arg to the Cys residues in the primary sequences of all vertebrate and invertebrate MTs suggests that the Lys and Arg residues also play a role in the formation of the metal complexes. In fact it has now been shown by a variety of techniques that, in the metallated forms of MTs, the protonated forms of the ε-amino groups in the Lys and Arg sidechains are markedly stabilized. This has been interpreted in terms of hydrogen-bonding interactions between the protonated amino groups and the negatively-charged metal-thiolate complexes, but its role in the metal-binding capacity of the MTs is unclear. However, a variety of spectroscopic techniques has shown that each bivalent metal ion is bound to 4 thiolate ligands arranged in tetrahedral symmetry. This requires sharing of some of the thiolate ligands by adjacent metal ions, in view of the ratio of about 3:1 for Cys residues to metal ions bound.

Even more detail has been provided by ^{113}Cd NMR experiments. Fully metallated MTs contain two entirely separate metal-thiolate clusters. In mammalian MTs cluster A contains 4 metal ions plus 11 Cys residues, while cluster B contains 3 metal ions plus 9 Cys residues. In crab MTs each of cluster A and cluster B is similar to cluster B of

the mammalian MTs. Chemical evidence for this partitioning into two well-separated clusters comes from the observation that, under appropriate conditions, mammalian MTs can be cleaved by subtilisin to yield similarly sized fragments. One of these fragments contains the C-terminus (residues 31-61), comprising 11 Cys and binding 4 metal ions, and corresponds to cluster A. The other fragment comprises residues 1-30 (N-terminus portion) containing 9 Cys and binding 3 metal ions. The most direct confirmation of the two-cluster model, however, has come from X-ray crystallographic determinations.

The potential importance of these cluster structures for the present work lies in the evidence that the stabilities of clusters A and B differ appreciably. Acidification studies have shown that metal ions are lost preferentially from cluster B. At pH 8.6 it was found that iodoacetamide reacts about three times faster with Cys residues in cluster B than with those in cluster A. Similarly, under conditions where most of the Cys residues are protonated (neutral pH or below), reconstitution of fully-metallated MT from apoMT plus metal ion proceeds *via* preferential formation of the entire cluster A, followed by cluster B. Cluster formation under these conditions is strongly cooperative, *i.e.* no non-clustered intermediates were observed. Thus, in any chemical modification of Cys residues in metallated MTs it must be kept in mind that two domains of very different reactivities are involved. That any reaction at all occurs is probably due to the dynamic fluctuations of the metal-thiolate bonds within each cluster, as shown by ^{113}Cd NMR studies. For example, the high mobility of Cd (II)

within cluster B is thought to underly the remarkably facile interprotein metal redistribution between $^{113}Cd_7$ - MT and Zn_7-MT. All of these chemical studies have involved mammalian MTs, but NMR studies [30] have suggested that cluster A in crab MT is again more stable (^{113}Cd more strongly bound) than cluster B.

Mass Spectrometric characterization of proteins.

Prior to 1981, mass spectrometry was of little or no use to protein and peptide chemists. In that year the fast atom bombardment (FAB) ionization technique was introduced [31]. This technique permits molecular weight determinations of peptides with molecular masses up to 3 kDa or so, on a fairly routine basis, and of proteins up to about 20kDa with fairly heroic efforts [32]. Development of techniques for tandem mass spectrometry has led to sequencing of peptides *via* interpretation of the fragment ion spectra of the MH^+ ions produced by FAB ionization of a peptide M [33]. Most of these developments involved analysis of purified peptides, introduced to the mass spectrometer *via* a batch probe. Subsequently, development of continuous-flow FAB ionization [34] permitted analysis of peptide mixtures by coupling the mass spectrometer directly to high performance liquid chromatography (HPLC).

More recently, however, biochemical applications of mass spectrometry have been revolutionized by the development of atmospheric pressure spray ionization

techniques. An excellent review of the development of these techniques and their biochemical applications has been published [35]. These techniques will all be referred to here as "electrospray", for convenience, although the pneumatically-assisted variant known as "ionspray" has appreciable advantages when coupled to HPLC [36].

The applications of electrospray ionization to protein characterization can be conveniently divided into three categories:

a accurate and precise ($\pm < 0.01$ %) determinations of molecular weights (up to about 200 kDa), by exploitation of the multiple-protonation phenomenon in which a protein M is transformed into a series of gaseous ions $(M + nH)^{n+}$, where the value of n can be as high as 10 for a given protein M; it is the high degree of redundancy in the resulting mass spectrum which permits the high precision in the (isotope-averaged) molecular weight measurement; specific derivatization reactions, *e.g.* for the thiol groups of Cys sidechains, can be used in conjunction with such molecular weight measurements to count the number of specified groups (*e.g.* thiols) in the protein molecule;

b protein mapping by HPLC-MS analysis of appropriate protein digests (*e.g.* those produced by trypsin, cyanogen bromide, *etc.*); such a two-dimensional map (both retention time and molecular weight information) provides a highly specific

descriptor of a protein;

c determination of primary sequences of peptides produced by specific digestion procedures; in particular, "tryptic peptides" produced by digestion with trypsin form abundant doubly-protonated species $(M + 2H)^{2+}$, which usually readily fragment to provide redundant and easily interpretable information on amino acid sequence.

In the above discussion, the use of HPLC coupled on-line to mass spectrometry has been emphasized. However, it is also possible to interface capillary electrophoresis (CE) to electrospray mass spectrometry, although application to peptides and proteins requires modifications to the CE methodology [37]. The great advantage of CE over HPLC is its very high separatory power (up to 10^6 theoretical plates), while its main disadvantage arises from the low injection volume (usually < 10 nL) which has consequences for detection limits and sensitivity in general. These consequences can be ameliorated by "sample stacking" techniques.

The objectives of the present work were to investigate the applicability of these modern mass spectrometry techniques, alone or in combination with HPLC and CE, to the characterization of metallothioneins and to initiate such studies of metalliothioneins from crab (*Scylla serrata*).

B DESCRIPTION OF THE EXPERIMENTAL METHODS

Chemicals. Except for crab (*Scylla serrata*) and trout metallothioneins which were supplied by IOS, all other proteins (rabbit liver metallothionein isoforms I and II, and horse liver mixed isoforms) were purchased from Sigma Chemical Co (St. Louis, MO) and used without purification. Acetic, formic and trifluoroacetic acids were obtained from BDH Chemicals Ltd (Poole, UK). Methanol and acetonitrile (HPLC Grade, Caledon Laboratories Ltd Georgetown, Ont.) and distilled and deionized (18 Mohm) water (Milli-Q water systems, Millipore Inc., Bedford, MA, USA) were used in the preparation of the samples and mobile phases. TPCK trypsin was obtained from Worthington Biochemical Co. (Freehold, NJ, USA). Dithiothreitol (DTT), 2-mercaptoethanol, and buffer solutions were purchased from Sigma Biochemicals (St. Louis, MO, USA). All mobile phase solvents contained 0.1% trifluoroacetic acid (TFA), and were filtered (0.45 μm) and degassed prior to use.

Disulfide bond reduction and alkylation of metallothioneins. Denaturation and unfolding of proteins was achieved by dissolving 1 mg of the crab metallothionein in 6 M guanidine HCl, 1M Tris/HCl pH 7.5. Reduction and alkylation of cysteine residues were achieved by adding to purified metallothionein a 50-fold molar excess of DTT per mole of protein disulfide bonds (18 cysteines were assumed to be present per mole). The solution was kept at 50°C for 1 h, cooled to room temperature, then treated with 3 mg iodoacetamide (two-fold excess over DTT), and the reaction allowed to proceed for 15 min at room temperature. The reduced and alkylated proteins were later purified

by HPLC using a C_{18} reverse phase column.

Tryptic digestion. The denatured and alkylated metallothionein was dissolved in 0.5 mL ammonium bicarbonate (pH 8.0) to which 0.5 mL of TPCK trypsin (1 mg/100 mL, 0.1 mM HCl) was added. The tryptic digestion was conducted at 37° C for 3 h, and stopped by freeze-drying the sample on a SpeedVac concentrator. The sample was redissolved in 1 mL of water containing 0.1% TFA.

HPLC (UV detection). A HP1090L liquid chromatograph (Hewlett Packard Co., Palo Alto, CA, USA) equipped with a binary DR5 solvent delivery system, a HP1040 diode array detector and a HP7994A data system, was used in all LC-UV analyses. The detector was set to 214 nm with full UV spectral acquisition on peak detection. Separations were achieved using a 2.1 mm ID x 25 cm HPLC column (Vydac Separation group, Hesperia, CA) packed with 5 µm Vydac 218TP, and operating at a flow rate of 200 µL/min. Unless otherwise specified separations were conducted using a linear gradient of 10% to 90% acetonitrile in water in 30 min, with 0.1% trifluoroacetic acid (TFA) as modifier.

CE (UV detection). An Applied Biosystems (Foster City, CA, U.S.A.) Model 270A capillary electrophoresis system was used for all CE-UV experiments. An untreated fused-silica capillary column (Polymicro Technologies, Phoenix, AZ, U.S.A.) of 90 cm (72 cm to detector) x 50 µm I.D. was used in all separations. Sample introduction was performed using hydrodynamic injection by applying a preset vacuum

(17 kPa) to the detector end of the capillary for 3 s. This injection resulted in the introduction of approximately 10 nL of sample solution into the capillary. Electrophoretic separations were carried out in 20 mM sodium citrate (pH 2). The (variable) UV detector wavelength was set to 200 nm. Data acquistion and handling were accomplished using a Hewlett Packard HP 3396A integrator linked to an MS-DOS microcomputer using the Chromperfect data acquistion-processing software package (Justice Innovations, Palo Alto, CA, U.S.A.).

Mass Spectrometry. All mass spectra were acquired using a SCIEX (Thornhill, Ont.) API/III triple quadrupole mass spectrometer equipped with an atmospheric pressure ionization (API) source operated in the ionspray mode. A MacIntosh IIx microcomputer was used for instrument control, data acquisition and data processing. A solution of horse heart myoglobin (1 mg/mL in 10% acetic acid) was used to tune the instrument for optimal sensitivity and to calibrate over the range m/z 600-2000. Ionspray mass spectra of metallothioneins were obtained by infusing a solution of 1 mg/mL of the proteins at a flow rate of 3-10 µL/min using a syringe pump (Harvard Apparatus, Southnatick, MA). Molecular weight determinations were made using at least 4 multiply charged ions from each multi-protonated peak series.

The LC-MS system consisted of a HP 1090L liquid chromatograph, used to generate gradient profiles at a flow rate of 200 µL/min, and coupled to the ionspray interface via a flow-splitter such that only 10-15 µL/min was introduced to the mass spectrometer. The voltage on the ionspray needle was maintained at approximately 5.6 kV. High-

purity air was used as nebulizing gas at an operating pressure of 90 psi (approximately 0.6 L/min). Separations of tryptic digests were obtained using the chromatographic system described above. A post-column Valco submicroliter injection valve with a 1 µL loop was used to optimize the LC-MS interface. LC-MS analyses were performed in either full mass scan mode or in selected-ion monitoring mode, using dwell times of 5 msec per Da or 200 msec per channel, respectively.

Combined LC-MS-MS analyses were achieved using the same chromatographic system as that described above. Tandem mass spectra were obtained following mass selection of precursor ions by the first quadrupole while scanning (typically 200-1500 Da) the third quadrupole mass analyzer with a dwell time of 5 msec per Da. Collisional activation at energies of 35 eV (laboratory frame of reference) was obtained by introducing Ar into the RF-only quadrupole at pressures sufficient to attenuate the precursor ion beam by 80-90%. This attenuation corresponded to a target gas thickness of typically 4×10^{14} molecules/cm^2.

The CE-MS experiments were conducted using techniques which were still under development at the time that the work on metallothioneins was performed. For this reason the details are given in the appropriate section, below.

C. ANALYTICAL METHODS WITHOUT MASS SPECTROMETRY

The quantitative determination of metallothioneins in various matrices has been accomplished using a variety of methods. These methods include standard techniques of biochemistry and immunochemistry [38]. Other methods include high performance liquid chromatography (HPLC) using detection by atomic absorption spectrometry [39], by inductively-coupled plasma emission spectroscopy [40], or by UV absorption at 214 nm [41]. Electrochemical quantitation of purified metallothionein extracts has also been described [42]. Very recently capillary electrophoresis (CE) with UV detection at 214 nm was demonstrated to be useful for qualitative and quantitative analysis of metallothionein extracts [43].

Examples of analyses of metallothionein preparations, by the LC/UV method described in the "Experimental Methods" section, are shown in Figure 1. Since the pH of the mobile phase was ≤ 2 (due to the trifluoroacetic acid), both the endogenous Cd^{2+} and Zn^{2+} would have been removed [6]. Each of the commercially available isoforms of rabbit liver metallothionein is clearly a complex mixture, reflecting the amino acid substitutions known to occur in rabbit metallothionein [6]. The horse metallothionein sample was a mix of isoforms, in any case. On the other hand, the crab metallothionein isoform I is much less complex (Figure 1c), with one major peak in the chromatogram accompanied by at least one partially resolved variant. Thus, LC/UV analyses do provide considerable information concerning the complexity of a

metallothionein preparation which may have been considered to be a single component based upon classical electrophoretic techniques.

Capillary electrophoresis (CE), however, can provide separations with extremely high resolving power. CE is a new approach to electrophoresis in which the traditional gel slabs, paper sheets, etc., are replaced by narrow bore (typically 50-100 μm i.d., and approx. 360 μm o.d.) fused silica capillaries of length 50-150 cm. The capillaries are coated on the outside with a thin polyimide film in order to preserve flexibility and mechanical strength. CE accommodates very high voltages (up to 30 kV, giving field strengths up to 600 V/cm) and current densities (up to 5A/cm^2, equivalent to currents of up to 300 μ A depending on the nature of the supporting buffer) because of the efficient dissipation of Joule heat made possible by the large ratio of surface area to volume. In turn this efficient cooling results in minimal radial temperature gradients, thus minimizing problems associated with convection and variations in viscosity across the capillary cross section. This radial uniformity is the ultimate guarantor of the very high separation efficiencies (up to 10^6 theoretical plates in 20-25 min) achievable using CE.

It is not appropriate here to describe in detail the physical principles underlying analytical CE; excellent expositions are available in the literature [44-46]. However, it is worthwhile for present purposes to emphasize that emergence of an analyte from one end of the capillary is the result of interplay of two different transport mechanisms. The first of these is the electrophoretic mobility of the charged analyte species through

the supporting buffer solution, under the influence of the applied field; this mobility is a function of the size and shape of the solvated species, and of its net charge. The second transport mechanism is electroendosmosis, which is the bulk flow of liquid resulting from the effect of the applied field on the electrical double layer adjacent to the capillary wall. Figure 2 illustrates these two effects as a function of pH, for the present case where the net charge on the silica internal wall is negative so that the adjacent double layer carries a net positive charge, and the annulus of positive charge is drawn towards the negative electrode. The bulk liquid flow that results is characterized by its flat profile, in contrast to the parabolic profile typical of viscous flow induced by pressure difference; thus the flow profile does not contribute to band broadening, as is the case for HPLC for example. The other feature of CE which will be important for Section D3 below is the very low volume flow rate (of the order of 100 nL/min) emerging from the capillary. Incidentally, when desired it is possible to reverse the direction of the electroendosmotic flow from that (Figure 2) dictated by the usual negative charge on the surface of the untreated quartz capillary. This reversal can be achieved [37] by coating the internal wall with a suitable compound which presents a positively-charged surface layer to the running buffer solution within the capillary. This reversal has been shown [37] to be beneficial for analysis of peptides and proteins by CE combined with mass spectrometry (Section D3).

A block diagram of a CE apparatus is shown in Figure 3. From this diagram it is clear that the electrophoresis column forms part of a simple DC electrical circuit, and that

questions of electrical continuity, at all points in this circuit, are crucial. This apparently trivial comment will assume great importance in Section D3, dealing with the coupling of CE to electrospray mass spectrometry. Injection techniques are also an important consideration in CE. Although it is not appropriate to enlarge upon this question here, it is worth commenting that the experimentally convenient approach of electrokinetic injection introduces a bias into relative quantities of components of a mixture, due to differences in mobilities. Thus, the electrokinetic injection technique is not suitable for quantitative analysis; hydrostatic and hydrodynamic injection techniques are much superior in this regard. Note also that, in conventional CE, the injection volume must be restricted to less than 2% of the total column volume (*i.e.* less than 20 nL or so for a CE column which is 50μm I.D. and 1m long), in order to ensure that the volume factor does not become the limiting determinant of CE peak width. Choice of running buffer is also crucial; it must be compatible with the detection used, *e.g.* transparent at the monitoring wavelength if a UV detector is to be employed.

An example of the application of the CE-UV technique, to analysis of complex metallothionein preparations from vertebrate species, is shown in Figure 4. Note that here again the pH was maintained at 2.0, guaranteeing complete demetallation [6]. The electropherogram of the total metallothionein (mixed isoforms) from rabbit liver, shown in Figure 4a, is a superposition of those for the pre-separated isoforms I and II (Figures 4b and 4c, respectively). The benefits of this pre-separation can be seen in

the degrees of complexity thus revealed for each of the two isoforms. This complexity confirms and expands upon that revealed by LC/UV analyses of the same two samples (Figures 1a and 1b). The electropherogram (Figure 4a) of the unseparated mixture of isoforms should be compared with Figure 2 of Ref. [43]. In the latter electropherogram the entire metallothionein mixture eluted from the capillary in a period of only one minute, so it is not surprising that a much lower degree of complexity was revealed [43] than in the present work (Figure 4). Note, however, that the present work used a sodium citrate buffer at pH 2.0, when metallothioneins are known [6] to be stripped of their metal content, while the work described elsewhere [43] employed an alkaline buffer (pH 9.1) which guaranteed full retention of the metal load.

As anticipated, therefore, the LC/UV and CE/UV techniques provide useful tools for establishing the complexity of metallothionein preparations. However, in order to investigate the nature of this complexity it is necessary to employ a more informative detection method together with the high-resolution separatory techniques. The only detector, currently available, which can provide additional chemical information for components emerging from a HPLC column or CE capillary, is mass spectrometry.

D. MASS SPECTRAL ANALYSES OF METALLOTHIONEINS USING ELECTROSPRAY IONIZATION.

In conventional mass spectrometry, using electron ionization (EI), chemical ionization (CI), or batch-mode fast-atom bombardment (FAB) ionization, the sample can be introduced *via* an appropriate direct insertion probe.

No such technique is possible for electrospray ionization, which is intrinsically limited to analytes in a flowing liquid solution. The injection technique, which plays the same role for electrospray ionization as the direct insertion probe does for EI, CI and FAB ionization, is flow injection analysis (FIA) whereby a small volume of analyte solution is injected into a continuous flow of solvent. This is achieved through use of the loop-injector technology developed for HPLC. As is the case for FAB ionization, the analyte ions can be readily suppressed by co-dissolved salts or other compounds. Accordingly, electrospray ionization using FIA is useful only for samples which have been properly cleaned up, particularly with respect to salts.

D1. Mass Spectrometry with Flow Injection.

The electrospray mass spectrum of a solution (pH 2.8) of the commercial preparation of rabbit liver metallothionein isoform I is shown in Figure 5. At this pH most of the metal ion content is likely to be displaced by protons [6]. Figure 5a shows an

overview of the entire mass spectrum, while Figure 5b shows a detailed expanded view of the raw spectrum over a limited m/z range. The spectrum is typical of electrospray mass spectra of proteins, in that the same pattern of peaks is repeated several times (though with varying intensities) over the m/z range. This is due to variations in the number of ionizing (external) protons bound to the protein molecule. By assuming that adjacent peak clusters differ only by one more or fewer ionizing proton, it is a simple matter to transform the mass spectrum from an m/z scale to a molecular weight scale. In this way all of the repeating clusters are superimposed, thus (at least in principle) yielding improved signal/noise ratio in the "molecular weight" spectrum, *e.g.* that for the rabbit liver metallothionein isoform I, shown in Figure 5c. The heterogeneity of this commercial preparation, demonstrated previously by LC/UV and CE/UV (Figures 1 and 4, respectively), is confirmed by Figure 5. Analogous data for rabbit isoform II are shown in Figure 6. Clearly the isoform II preparation is contaminated by isoform I, as demonstrated by the prominence in Figure 6c of peaks corresponding to molecular weights 6145, 6175(7), 6214 and 6240 Da. This contamination is also evident in the chromatograms and electropherograms of Figures 1 and 4.

Some indication of the long-term reproducibility of these measurements is given by a comparison of Figures 6 and 7c, each of which shows electrospray ionization mass spectra of rabbit liver metallothionein (isoform II). The same commercial preparation was used in both cases, but the protein solutions were made up fresh on each

occasion (separated by several months). Thus, the pH of the solutions, as well as other operating parameters, probably differed slightly. Indeed the relative intensities of the peak clusters corresponding to different degrees of protonation do differ, but the two spectra are in reasonable agreement when transformed to the molecular weight scale (Figure 6c and insert on Figure 7c).

When the pH of the solutions was increased to 4.5, the appearance of the electrospray mass spectra changed dramatically. Examples were shown in a preliminary account of the present work [47], and are not reproduced here. As the pH was increased the signal/noise ratio in the spectra dropped dramatically, the degree of protonation dropped (*e.g.* from a maximum of 9+ at pH 2.8 to only 7+ at pH 4.5), and the deduced molecular weights rose appreciably. The data thus obtained, for rabbit metallothionein isoforms I and II, are compared in Table 2.

It is likely that these dramatic differences reflect varying degrees of displacement of the endogenous metal ions as the pH is varied. Zinc ions are reported to be displaced completely at pH3, while cadmium ions are completely lost at pH2 [6]. Recently [48] an independent study of rabbit liver metallothionein isoform II, purchased from the same commercial supplier, was published. The authors of this study obtained the metal-free protein (apometallothionein) by treatment with 0.1% aqueous trifluoroacetic acid, followed by gel filtration. The resulting metal-free protein was a complex mixture, and a cut from a purification procedure using reversed-phase HPLC

was analyzed by electrospray mass spectrometry. Even this highly purified fraction contained two components, with molecular weights 6126 and 6156 Da [48]. The former value correponds to that calculated from the published sequence for rabbit liver metallothionein 2a [49] (see Table 1), if the N-terminus blocking group is assumed to be acetyl. The higher value does not correspond to any published sequence, but is consistent with replacement of one A (Ala) residue by T (Thr).

At this point, it is appropriate to attempt to draw a few conclusions. It is clear that modern high-resolution separatory techniques have exposed the complexity of metallothionein isoforms prepared "pure" by classical biochemical techniques. This is exemplified by the present work (Figures 1 and 4-7) and by the published electrospray mass spectra [48] of rabbit liver isoform IIa, thus shown to be comprised of two components (labelled 2a and 2a' [48]) differing by 30 Da. It is clear that published amino acid sequences of preparations of rabbit metallothionein isoforms, which have not been subjected to meticulous purification, should be treated with some reserve. The other conclusion to be drawn from a comparison of the results of the present work on electrospray mass spectrometry of metallothionein preparations (*e.g.* as summarized in Table 2), with that of the published data [48], is that it is wise to first prepare the demetallated apometallothioneins off-line. Fractionation by semi-preparative HPLC would also be advantageous if directly-coupled LC/MS techniques are not to be used. Present attempts to prepare the apoproteins by displacing the metal ions by protons, in a single solution, led to difficulties of interpretation. For

example, it is not clear whether the various species recorded at pH 2.8 in Table 2 represent heterogeneity of the apometallothioneins or variations in residual metal content (but see the LC/MS section below). Successful demetallation and purification procedures are described by Yu *et. al.* [48].

D2. LC-MS Analysis.

The results of LC-MS analysis of a commercial preparation of rabbit liver metallothionein isoform I are summarized in Figures 8 and 9. The experimental conditions used are described in the Experimental section. Note that the use of 0.1% trifluoroacetic acid in the LC mobile phase implies that the pH was close to 2, the condition at which the protein is believed to be completely demetallated. Figure 8a shows the same LC-UV chromatogram as was shown in Figure 1a, for comparison with the LC-MS data which were obtained in the same experiment which used the UV and MS detectors in series. The displacement of the LC-MS retention times (Figure 8b) by about 2 minutes reflects the flow time through the capillary connecting the UV detector to the electrospray needle.

The LC-MS experiment was conducted in selected ion monitoring (SIM) mode. The eight m/z values monitored were chosen to be representative of the four components of this preparation, detected and characterized previously at pH 2.8 by flow-injection electrospray mass spectrometry (Table 2). The eight m/z values monitored, and the

molecular weights to which they correspond, were 878 and 1025 (6144), 883 and 1030 (6174), 888 and 1036 (6214) and 892 and 1041 (6241). Figures 8c and 8d indicate that the earliest eluting component is that with molecular weight 6144 Da. A more complete representation of these LC-MS data is shown in Figure 9. The three major components are those of molecular weights 6144, 6214 and 6241 Da, while the minor component of molecular weight 6174 Da was not resolved from that at 6144 Da under the chromatographic conditions used.

In the context of the conclusions drawn in Section D.1 above, it is important to note that the four species characterized by flow-injection electrospray mass spectrometry at pH 2.8 (Table 2) are also chromatographically distinct species (Figures 8 and 9) at pH 2.0 or so. Thus, it seems likely that these observations correspond to heterogeneity of the amino acid composition, rather than to poorly defined quantities of residual Cd^{2+} and/or Zn^{2+}. This is also confirmed by the complexity observed in the CE-UV electropherogram (Figure 4b) of this same preparation, which was obtained using a buffer of pH 2.0. It thus seems worthwhile to attempt an interpretation of these data, in the context of the amino acid sequence published for rabbit liver metallothionein isoform I.

The two species known [6] to be present within isoform I differ at residue position 54 as defined in Table 1, where Leu and Ser can occupy this position (113 *vs* 87 Da, a difference of 26 Da). Note that the four species detected by electrospray MS and LC-

MS (see Table 2, pH 2.8) fall into two pairs, 6144 and 6174 Da (difference of 30 Da but the second of these is an ill-defined minor component), and 6214 and 6241 Da (difference of 27 Da). Thus, a reasonable hypothesis would be that, within each pair characterized in the present work, the mass difference is due to this Leu/Ser heterogeneity. However, it is still necessary to account for the molecular weight difference between the two pairs of proteins (about 170 Da), and for the difference between the molecular weights measured here and those calculated from the published sequences (6043 and 6069 Da, Table 1). If the two published sequences are also considered as a Leu/Ser pair, their molecular weights are 101-105 Da lower than the first pair measured here (6144 and 6174 Da), and 171-172 Da lower than the second pair (6214 and 6241 Da). Obviously the two pairs characterized in the present work differ by some 67-70 Da with the higher value regarded as the more reliable (since the 6174 Da molecular weight is the more uncertain).

Before speculating on the amino acid differences which could account for these observations, it is worthwhile to note some background information. The isotope-averaged atomic weights of Cd and Zn are 112.4 and 65.4 Da, respectively. The value for Zn is uncomfortably close to the molecular weight difference (67-70 Da) between the two pairs observed in the present work (Table 2, pH 2.8). However, Cd^{2+} is appreciably more strongly bound than Zn^{2+} (see Introduction), so any residual metal ion is unlikely to be Zn^{2+}. Moreover, metal ion binding by metallothioneins in acidic solution is strongly cooperative in view of the thiolate-metal cluster structures, and

while either one or both clusters can be formed no intermediate metal-protein ratios have been observed at low pH. Finally, for the isoform II preparation from the same supplier, Yu *et al.* [48] found that Cd^{2+} dominated Zn^{2+} by a ratio of about 5:1. Thus it is deemed improbable that the 67-70 Da difference (70 Da being the more reliable estimate), between the two pairs observed (Table 2) for isoform I, can be attributed to a single residual Zn^{2+} ion.

The discrepancy between the molecular weights measured in the present work, and those calculated from published sequences, are more troubling. As emphasized by Hunziker [49], direct N-terminal sequence analysis (Edman) of verterbrate metallothioneins is impossible due to the acetylation of their N-terminal Met residue. However the Asp^2 - Pro^3 bond is readily cleaved by acid, and this fact has been exploited to unblock the N-terminus and thereby permit reliable sequencing up to about residue 31. The same problem exists for the peptide containing the original N-terminus, formed as a hydrolysis product by specific enzyme action (*e.g.* trypsin). There is therefore an inherent uncertainty in the N-terminus sequence, which Hunziker [49] states to have been determined by amino acid analysis of the fragment released by acid cleavage of the Asp -Pro linkage. This is actually not a sequence determination at all, and it is at least possible that an additional amino acid residue would be missed in such a procedure, or possibly that the N-terminus blocking group is not acetyl but something considerably larger. The major part of each of these sequences has, of course, been confirmed by redundant information on sequences of

peptides formed by various specific enzyme digestions.

Possible interpretations of the difference of 101 Da include the presence of an additional Thr residue (101 Da) not detected in the sequencing of the N-terminal blocked protein, as discussed above, or possibly replacement of an acetyl group (43 Da) by $C_8H_{17}C0$ (141 Da, a difference of 98 Da). Such odd-numbered fatty acid residues are not common, however, and this latter possibility seems much less likely. Similarly, the additional 70 Da or so, separating the two pairs of compounds characterized in the present work, could be accounted for by an extra Ala residue (71 Da) in addition to the Thr residue already suggested. Both Ala and Thr residues are fairly abundant in vertebrate metallothioneins (Table 1).

While these interpretations are necessarily speculative at this point, the present work has strongly suggested that the primary sequences of these metallothioneins need to be re-examined. One reason for this is the N-terminus problem discussed above. The other is the complex mixture status (Figures 1 and 4-7) of most of the isoform preparations which have been subjected to conventional sequencing procedures.

D.3 CE-MS Analyses.

Since the LC-UV and CE-UV analyses had already confirmed that better separations can be obtained by using CE, it was decided to conduct mass spectrometric analyses

of crab metallothionein, the protein of interest, by CE-MS. Before proceeding to describe the results obtained, it is appropriate to briefly describe the experimental procedures used, since these are not widely known.

An apparently trivial requirement for successful CE-MS operation is that the CE instrument can be located reasonably close to the mass spectrometer, to minimize peak broadening associated with a long transfer line. The Applied Biosystems Model 270A CE System, used in the CE-UV experiments, was replaced in the CE-MS work by a more recently acquired Beckman P/ACE 2050 model since the latter has pressure injection capabilities while the former relies on electrokinetic injection. The CE-MS coupling was achieved using a modified fully-articulated Ionspray® interface and a coaxial capillary arrangement of design similar to that described by Smith *et. al.* [50,51]. Details of the construction of this CE-MS interface have been presented earlier [52] and only a brief description will be presented here. In this configuration (Figure 10) the back tee is used to provide the make-up solution (25% methanol, 0.2% formic acid) necessary for the maintenance of good electrical continuity during the CE separation. This capillary arrangement minimizes column peak broadening and ensures good electrical contact between the electrospray needle and the CE electrolyte. The make-up solution is delivered to the back tee by a syringe pump (Applied Biosystems, Foster City, CA, U.S.A.) at a flow rate of 5-10 μL min. A submicroliter injector (typically 0.1 μL loop) can be incorporated at this point, not only to provide a means of optimizing the ionspray conditions prior to CE/MS analysis, but also to permit calibration of the CE injection system for quantitative analyses by

injecting a suitable standard. The make-up solution is introduced directly into the ionspray needle, and the front tee is used only to introduce the nebulizing gas. The CE column protrudes *ca.* 0.1 mm from the ionspray needle which in turn protrudes *ca.* 0.5 mm from the nebulizer sheath to produce a spray of charged droplets. Although different sizes of capillaries and stainless steel ionspray needles have been used at IMB, best performance and easiest optimization have been achieved using small columns (*e.g.* 23-gauge stainless steel tubing for the needle, and 20-50 µm I.D., 180 µm O.D. fused silica capillaries)

CE-MS separations were carried out using a 1.1 m x 50 µm I.D. fused silica capillary (Polymicro Technologies). Prior to each separation the capillary was coated with an aqueous solution of 5% (w:v) hexadimethrine bromide, (Polybrene®, Aldrich Chemicals, Milwaukee, WI) and 2% (w:v) of ethylene glycol. The reasons for this coating procedure are described in detail by Thibault *et. al.* [37], and are briefly outlined here. These reasons all derive from the fact that the inside wall of untreated fused quartz capillaries bear a negative net charge. If the pH of the working buffer is such that the proteins to be analyzed bear a net positive charge, the electrostatic attraction from the quartz walls is sufficient to greatly degrade the efficiencies of both separation and detection. If the proteins are arranged to have a negative charge by using a buffer with high pH, very poor MS sensitivity is obtained in positive ion mode. It is believed [37] that this is due to a combination of residual wall adsorption effects, to contributions from the CE buffer to the total electrospray ion current (and thus suppression of the desired protein signals), and to problems in converting anionic

species emerging from the CE capillary into their cationic counterparts by the electrospray process. In addition, under such high pH conditions in untreated capillaries, the anionic proteins have electrophoretic mobilities in the direction opposite to that of the electrosmotic flow, which is the overwhelming influence in directing the analytes towards the detector. By coating the capillary walls with a cationic substance (several proprietary preparations are available), it is possible to use acidic buffers. Then the proteins are present as cations, and their electrophoretic motion opposes that of the electroosmotic flow which is now due to a negatively charged sheath of anions adjacent to the positively charged capillary wall (compare the more usual situation in Figure 2). In order for the analyte proteins to migrate towards the detector it is now necessary to reverse the polarity across the capillary. This approach was shown [37] to provide excellent efficiencies of separation and detection for proteins in acidic buffers.

In the present work, unless otherwise indicated, all separations were performed using a 15% acetic acid buffer and an effective (net) potential drop of 30 kV across the capillary. Hydrodynamic injections, on the Beckman P/ACE Model 250 CE instrument, were performed by pressurizing the sample vial at 5 p.s.i. for 20s while the end of the capillary was immersed in the sample solution.

Optimum CE separations (peaks no more than 10s wide) imply that the mass spectrometer scan cycle should be no more than about 2s if sufficient scans are to be

obtained across a CE peak. Some compromise between spectral distortion (slow scans) and poor signal levels (faster scans) must be sought. In the particular case of the SCIEX API III mass spectrometer used in the present work, a stepping scan function tied to ion-counting detection is used. Thus, the compromise to be sought on scan speed actually involves two selectable parameters, the step size (usually in the range 0.1 to 1 m/z unit) and the time per step (usually in the range 5 to 200 ms).

Application of this CE-MS technique to isoform I of metallothionein from crab (*Scylla serrata*) gave the results summarized in Figure 11. The major component eluting at 14.7 min, with a width at half-height of about 25s, is accompanied by low levels of poorly resolved components. The mass spectrum shown in Figure 11, plotted on both the raw m/z scale and also transformed to the molecular weight scale, indicates that the major component has a molecular weight of 6060 Da. Minor components of molecular weights 5771, 5886, 6070 and 6080 Da, are also apparent although the uncertainties in these molecular weights are appreciably greater than for that of the main component. This information is to be compared with the value of 6001 Da calculated from the published sequence (Table 1).

The discrepancy of 59 Da is unlikely to be accounted for by attachment of one Zn^{2+} ion (65.4 Da), for the same reasons as those discussed in the context of the LC-MS data for rabbit metallothionein (Section D.2 above). The N-terminus of crab metallothionein is <u>not</u> blocked, unlike that for most vertebrates, so possible

uncertainties arising from the N-terminal residues are not relevant in this case. The discrepancy might correspond to an additional Gly residue (57 Da) which could have been somehow missed in the original sequence determination [29]. Alternatively, one or more residues might have been misassigned. Any further speculation is unwarranted at this point. However, as in the case of the rabbit liver metallothionein, the mass spectrometry data have cast serious doubt on the total accuracy of the published sequences. The molecular weight discrepancies are larger than the experimental uncertainties by more than one order of magnitude.

E. PRIMARY STRUCTURE CHARACTERIZATION OF METALLOTHIONEIN FROM CRAB *(SCYLLA SERRATA)*

In view of the doubts raised by the present work concerning the primary sequences of metallothioneins (for a proven case where mass spectrometry corrected a published sequence of an unrelated protein see [53]), it was decided to investigate the particular case of principal interest in the present work (*Scylla serrata*) by mass spectrometric methods, and to compare the results for comparison with the conclusions reached using classical biochemical sequencing techniques. This particular case provides a good example for a first attempt at such a comparison, since crab metallothionein isoform I appears to comprise just one major constituent with only minor contributions from variants (Figure 11).

The general strategies used in this part of the work are the same as those followed in a previous investigation of primary sequences of cysteine proteases isolated from lobster [54]. Thus, attempts were made to count the number of cysteine residues by comparing the molecular weight of the native protein with that measured following reduction of any disulfide bridges and specific alkylation of the Cys residues. Subsequently, tryptic digests of the reduced and alkylated proteins were analyzed by LC-MS to provide a tryptic map, and by LC-MS/MS to obtain sequence information for the tryptic fragment peptides. It was decided to use LC rather than CE for these experiments, since sample loadings permissible for conventional CE are too low to provide adequate signal/noise ratios in MS/MS spectra obtained in on-line experiments. Sample-stacking techniques, whereby up to 50% of the CE capillary can be filled with the test solution with only minor consequences for the separation efficiency, are under development at IMB.

E1. Determination of number of cysteine residues.

The strategy here is to denature the protein sufficiently that all cysteine residues are exposed to the alkylating agent. This is usually achieved by dissolving the protein in a concentrated solution of guanidine hydrochloride (6 molar) to break all intramolecular hydrogen bonds. Reduction of any disulfide bonds is accomplished by adding a large excess (50 times the estimated molar concentration of Cys residues) of dithiothreitol (DTT). Usually the solution is maintained at a slightly alkaline pH (7.5-8.0), in the

present case by using a 1 molar Tris hydrochloride buffer. The time required for the denaturation/reduction process varies from case to case, and can only be determined by trial and error. In the present case the mixture was maintained at 50°C for 1 hour. Subsequent alkylation of those reduced Cys residues which are accessible to the ambient solution can be achieved in a variety of ways. The alkylating reagent used in the present work was iodoacetamide, which reacts with the Cys thiol group by replacing the thiol hydrogen by a ($-CH_2$ - C0 - NH_2) group, *i.e.* an increase of 57 Da per thiol group thus derivatized. In practice the denatured solution was cooled to room temperature, and then treated with a quantity of iodoacetamide corresponding to a two-fold molar excess over the DTT added previously. Again the reaction time required is somewhat dependent on the particular protein, but usually 15-30 min suffice. This alkylation procedure was one of those used in the original investigation [29] of the primary sequence of crab metallothionein. However it should be noted that S-methylation has been recommended for use in Cys modification of metallothioneins [55]. This procedure is infrequently used since the methylated products tend to be insoluble in water, but this difficulty is not encountered in the case of metallothioneins [55]. However, for present purposes of attempting to count Cys residues *via* molecular weight shifts, as a check on amino acid analysis, the larger shifts obtained using iodoacetamide are beneficial.

The results obtained by applying the procedures described above to crab metallothionein isoform I, are summarized in Figure 12. Analysis of the reaction

mixture by LC-MS, rather than by flow-injection electrospray MS resulted in separation of the modified proteins from the guanidine hydrochloride and Tris buffer which were believed to have been detrimental to the spectral quality obtained previously [54] in the case of the lobster cysteine proteinases. The total ion chromatogram (TIC) shows two well-separated peaks at 22.4 and 23.0 min. The mass spectra obtained at each chromatographic peak are shown in Figure 12, on both the m/z and transformed moleuclar weight scales, and are indeed of much better quality than those obtained previously [54] using flow injection of the entire reaction mixture. The two major components have molecular weights of 6636 and 6580 Da, a difference of 56 Da corresponding to one extra alkylation (the more fully alkylated species elutes first under reversed-phase conditions). However minor quantities of other components, with molecular weight spacings of approximately 57 Da, are also observed (Figure 12).

Interpretation of these results must take into account the fact that the metallothionein sample subjected to reduction and alkylation procedures was <u>not</u> the demetallated apoprotein. Since the native metal ions were still bound at the alkaline pH used, the alkylating agent had to displace the metal in order to access the thiolate groups. In this regard the role of the reducing agent (DTT) was probably restricted in this case to preventing any side-reactions forming disulfide bridges. It is interesting that one of the major components (molecular weight 6636 Da) corresponds to a shift of 576 Da from the non-alkylated protein *i.e.* very close to the value predicted from alkylation of 10 Cys residues, while the other major component (molecular weight 6580 Da)

corresponds to just 9 alkylated Cys residues. (Recall that the LC mobile phase had a pH of about 2, due to the TFA modifier, so that residual metal ions were almost certainly removed during the LC-MS analysis).

Crab metallothionein contains 18 Cys residues [6, 29], which form two stable metal ion - thiolate clusters each containing 9 Cys residues, as discussed in the Introduction. The less stable cluster B (see Introduction) is presumably attacked preferentially. It has been noted [6] that, at pH values above 7 and at metal/protein molar ratios below the full value of 7, the formation of isolated metal ion - thiolate complexes is favoured over that of clustered structures. Thus, it can be speculated that, once cluster B has been disrupted the metallothionein is no longer as stable under alkaline conditions and is more similar to other proteins in this regard. The fact that incomplete alkylation of the full 18 Cys residues (maximum of 12 observed in Figure 12; m/z 6749) was achieved is similar to the result obtained for lobster cysteine proteinases [54], even under conditions which led to the ideal (complete) alkylation result for hen egg-white lysozyme.

This observation emphasizes the dependence of the efficiency of this alkylation procedure on the particular protein under study. In view of the unique metal-binding capacity of metallothioneins, it is probably advisable in future work to first demetallate the proteins [47] and work with the resulting apometallothioneins.

E2. Tryptic digestion of the alkylated metallothionein.

The two major products of the alkylation of crab metallothionein isoform I (Figure 12) were collected using preparative HPLC. Each fraction was subjected to tryptic digestion as described in the Experimental section, and the resulting mixtures of tryptic fragment peptides were analyzed by LC-MS. Figure 13 shows the TIC thus obtained for the tryptic digest of the alkylated component of molecular weight 6580 Da, which eluted at 23.0 min (Figure 12). A two-dimensional contour plot of ion intensity (represented by depth of shading) as a function of both m/z and of retention time, is shown for the same LC-MS run in Figure 14.

For purposes of the following discussion, it should be appreciated that trypsin specifically hydrolyzes peptide bonds on the C-terminus side of the basic residues Arg (R) and Lys Typical tryptic fragment peptides have a strongly basic residue at the C-terminus, plus the basic N-terminal amino group, and thus readily form doubly-protonated species under electrospray conditions. The exception to this rule is the tryptic peptide containing the original protein C-terminus, which has no basic residue. Incomplete digestion by trypsin can yield peptides with one basic residue at the C-terminus plus another in the chain somewhere. Such peptides readily form both doubly and triply protonated species under electrospray conditions.

Mass spectra, obtained at the crests of the LC peaks shown in Figures 13 and 14, are

shown in Figures 15 and 16. (These spectra represent cross sections through Figure 14, parallel to the m/z axis). The large peak at 4.4 min (Figure 13) represents the void volume of the column, and the mass spectrum obtained at this time corresponds to all the buffer, salts, etc., present in the digestion media (see Figure 14 at 4.4 min.). The mass spectrum obtained at 6.4 min (Figure 15) contains two peaks, at m/z 492 and 983, corresponding to the $(M + 2H)^{2+}$ and $(M + H)^{+}$ ions from a peptide M of molecular mass 982 Da. The mass spectrum obtained at 12.5 min is similarly interpreted in terms of a peptide of molecular mass 965 Da. The intensities of the doubly-protonated species, in both cases, suggest that these peptides contain a basic residue (Arg or Lys) as expected for a tryptic fragment peptide. However, the spectra obtained at 11.5 and 4.9 min (Figure 15) are more problematic. The peaks observed at m/z 732 and 796 (11.5 min spectrum) could be either singly or doubly protonated species, whose doubly or singly protonated counterparts, respectively, would not have been observed due to the limited scan range (m/z 400 to 1400) used in order to maintain a suitably short scan cycle time. In fact the MS/MS spectrum of m/z 732 (see below) showed that this is a doubly-protonated form of a peptide of molecular mass 1462 Da. Similarly, the peak at m/z 671 in the spectrum obtained at 4.9 min (Figure 15) was shown by its MS/MS spectrum to be the MH^{+} ion of a peptide of molecular mass 670 Da.

The mass spectra shown in Figure 16 are more complicated. That obtained at 13.3 min (very weak peak in Figure 13) clearly corresponds to several species. The peaks

at m/z 1274 and 851 are the $(M + 2H)^{2+}$ and $(M + 3H)^{3+}$ ions of a peptide M of molecular mass 2549 Da, while the peaks at m/z 864 and 1295 are similarly interpreted in terms of a peptide of molecular mass 2589 Da. The fact that both doubly and triply protonated forms are observed suggests that these peptides contain at least two basic residues, and are thus possibly the products of incomplete tryptic digestion, although other explanations based on disulfide bridge formation are possible. The additional peaks at m/z 724, 732, and 797, in the mass spectrum obtained at 13.3 min, could be either singly or doubly protonated species for the same reasons as those discussed above for the uncertainty in the interpretation of the peak at m/z 732 (Figure 15). The mass spectra obtained at 15.7 and 16.3 min correspond to two poorly resolved peptides. The more abundant peptide has molecular mass 2532 Da, giving rise to the $(M + 3H)^{3+}$ and $(M + 2H)^{2+}$ ions at m/z 845 and 1267. Again, this peptide contains at least two basic residues, and thus possibly represents a product of incomplete tryptic digestion and/or of disulfide bridge formation. The considerably less abundant peptide has molecular mass 2586 Da, indicated by the ions at m/z 863 and 1294.

It is worthwhile at this point to summarize the results of this LC-MS analysis. Chromatographically resolved peptides, each containing at least one basic residue, eluted at 6.4, 11.5 and 12.5 min and had molecular masses 982, 1462 and 965 Da, respectively. A co-eluting peptide at 11.5 min is indicated by an ion at m/z 796, which could correspond to a molecular mass of either 795 or 1590 Da. The intense LC

peaks at 15.7 - 16.3 min (Figure 13) contain one abundant and one minor peptide, of molecular masses 2532 and 2587 Da, respectively, each of which contains at least two basic residues. The fairly intense LC peak, eluting just after the void volume at 4.9 min, corresponds to a singly-charged species at m/z 671. Finally, the weak LC peak at 13.3 min corresponds to several minor components including peptides containing at least two basic residues and of molecular masses 2549 and 2589 Da.

The published sequence [29] of crab metallothionein isoform I (Table 1) predicts that ten tryptic fragment peptides should be formed, eight of which have Lys (K) and one of which has Arg (R) at the C-terminus, together with the original C-terminal fragment (residues 52-58) of the protein. The proprietary software package "MacProMass" accepts a defined sequence and then calculates, amongst other things, the molecular masses of the tryptic fragments. Figures 17 and 18 show respectively the results of this calculation for the published sequence [29] and for the extreme case in which all Cys residues have been alkylated by iodoacetamide (57 Da increments). The particular protein subjected to tryptic digestion, and subsequently analyzed by LC-MS (Figures 13-16), had only 9 of the 18 Cys residues alkylated, with minor contributions from other degrees of alkylation. In this regard, it should be noted that the tryptic digestion was conducted at pH 7.5, so that disulfide bridge formation prior to trypsin hydrolysis represents a potential complication. In addition, the molecular mass of the underivatized protein was some 58 Da higher than that predicted from the published sequence. These complications make it extremely difficult to provide a complete

interpretation of the LC-MS data for the tryptic digest.

Accordingly, an extended version (not shown) of Figure 17 was created by hand calculation. In this extended version each of the tryptic peptide molecular masses was increased by increments of 57 Da up to a maximum specified by the number of Cys residues in that peptide. One additional increment of 57 Da was also included, on the simplifying assumption (made on the basis of no evidence) that the discrepancy between the measured molecular mass and that calculated from the published sequence corresponds to one extra Gly residue. (If this is not the correct interpretation of the mass discrepancy, it must correspond to replacement of one or more residues in the published sequence by other residues of higher molecular mass). This extended version also included predicted masses of tryptic hydrolysis products containing 3 and 4 basic residues, together with their partially or wholly alkylated counterparts. Matches were then sought for the observed molecular masses (Figures 15 and 16) within the extended list of predicted masses for the tryptic peptides.

Some good matches were obtained. The peptide eluting at 6.4 min with molecular mass 982 Da matches well the tryptic fragment (2,3), containing residues 9-17 inclusive (and thus two Lys residues), and with just one of the Cys residues alkylated. Similarly, the peptide of molecular mass 1462 Da eluting at 11.5 min corresponds closely to the predicted tryptic fragment (5,6) containing residues 27-38, including two Lys residues and with all 4 of the Cys residues alkylated. No match could be found

for the low-abundance compound of molecular mass 965 Da which eluted at 12.5 min. The other peptide eluting at 11.5 min, of molecular mass 795 Da, is reasonably matched by the prediction of 797 Da for the C-terminal tryptic fragment 10 (residues 52-58) with 2 of the 3 Cys residues alkylated. The two peptides eluting at 15.7-16.3 min (2532 and 2587 Da) are separated by 55 Da suggesting that they are related by virtue of one additional alkylated Cys residue, but no convincing matches with the predicted molecular masses could be found. This is particularly disturbing in the case of the 2532 Da peptide, which is a major product of the tryptic digestion (Figure 13). Finally, no match for the species at 791 Da eluting at 4.9 min, nor for the two peptides eluting at 13.3 min, could be found.

In view of the limited purity of the peptide HPLC fraction subjected to tryptic hydrolysis, failure to understand the low-abundance peptide masses should not be regarded as too serious. It is also worth mentioning in this context that trypsin provides self-hydrolysis products at low levels [56], which can confuse the interpretation of an LC-MS tryptic map such as that summarized in Figures 13-16.

However, the failure to match the major hydrolysis product of mass 2532 Da is a major problem. This peptide clearly represents a product of incomplete hydrolysis. The closest match is the predicted tryptic fragment (4,5,6) containing residues 18-38 and with 6 of the 7 Cys residues alkylated (calculated molecular mass 2540 Da). This discrepancy suggests that either the 58 Da difference, between the measured

molecular mass of the native protein and that calculated for the published sequence, is not in fact due to a single simple cause such as an additional Gly residue, or that this tryptic fragment peptide arose from hydrolysis of a protein with disulfide bridges.

E3. Sequence determination by LC-MS/MS

Tryptic fragment peptides, by virtue of the basic residue at the C-terminus and the comparably basic free primary amino group at the N-terminus, form abundant doubly-protonated species under the conditions of electrospray ionization. Upon low energy collisional activation (collision energies of a few tens of eV, characteristic of quadrupole instruments), these doubly protonated species generally dissociate to yield an intense fragment ion spectrum containing extensive information about the primary sequence. Such spectra are usually characterized by a complete (or almost so) series of **y"** fragment ions, accompanied by a less intense (and sometimes incomplete) series of **b** and/or **a** fragment ions. (The fragment ion nomenclature used here is that introduced by Roepstorff and Fohlman [57]). The mechanisms underlying this extremely useful property of tryptic fragment peptides have been examined [58]. In the present work, some of the tryptic peptides formed from S-alkylated crab metallothionein isoform **I** were examined in this way, to investigate their primary sequences.

An LC-MS/MS experiment was set up so that the precursor ions, selected by the first

quadrupole mass filter of the triple quadrupole mass spectrometer, were appropriate to the retention time windows defined in the LC-MS experiment (Figures 13 and 14). Only the more abundant peptides were examined in this way. The experimental conditions used are described in the Experimental section.

The fragment ion spectrum of the $(M + 2H)^{2+}$ ion (m/z 492), formed from the peptide eluting at 6.4 min, is shown in Figure 19b. Production of fragment ions at m/z values higher than that of the slected precursor ion is consistent with the interpretation of this m/z 492 ion as a doubly-charged species. A tentative assignment of this peptide as tryptic fragment (2,3), containing residues 9-17 with one of the three cysteine residues alkylated, was made in section E.2 above. The low-mass end (below m/z 160 or so) of a fragment ion spectrum like Figure 19b frequently contains information concerning the amino acid composition of the peptide [59]. The most important species in this context are the immonium ions, $H_2N^+ = CHR$, where R is the sidechain of the amino acid residue (-HN-CHR-CO-), so the molecular mass of the immonium ion is 27 Da less than that of the residue. In some cases other characteristic ions accompany or replace the immonium ions [59], *e.g.* for Arg the immonium ions (predicted m/z 129) are not observed, but instead a series of low-mass fragments at m/z 70, 87 and 100 indicate the presence of Arg residues. (If Arg is the C-terminal residue, as is the case for many tryptic fragment peptides, an intense $\mathbf{y_1}$" ion at m/z 175 is also an indicator for this residue).

The low-mass region of Figure 19b is dominated by m/z 133, which is that calculated for the immonium ion of a Cys residue which has been alkylated by iodoacetamide, *i.e.* the Cys residue [- NH - CH (CH_2 SH) - CO-] has molecular mass 103 Da, or 160 Da when alkylated (57 Da increment), so the predicted immonium ion should appear at m/z (160-27=133). Thus already we have good confirmatory evidence that this peptide contains at least one Cys residue alkylated to give (carboxyamidomethyl-Cys), which will be denoted by Cys*. The other intense peaks at m/z 101 and 84 correspond to the immonium ion of Lys and to subsequent loss of NH_3 from the sidechain and that at m/z 72 to the Val immonium ion. The predicted sequence (Figure 17) of the (2,3) tryptic peptide is CVCKEGGCK, so that this information from the low-mass region is consistent with the tentative identification.

The presence of two Lys (K) residues in this peptide, one at about mid-chain plus one at the C-terminus, means that the simple fragmentation mechanisms for doubly-protonated tryptic peptides [58] (K at C-terminus only) are disrupted by the tendency of both external (ionizing) protons to remain anchored on the two basic Lys sidechains, rather than mobilizing along the peptide backbone to initiate charge-site induced fragmentations. Thus, Figure 19 is <u>not</u> an example of the richly informative fragment spectra characteristic of true tryptic peptides. However, the spectrum does provide some useful confirmatory information.

The intense fragment ion at m/z 203 corresponds to the predicted $\mathbf{b_2}$ fragment (CV),

rather than to its alkylated version (C*V) predicted at m/z 260. This strongly suggests that the site of alkylation is not the C-terminal Cys residue. The other intense fragment at m/z 307 corresponds to the C-terminal fragment ion $\mathbf{y''_3}$ (GCK), whereas its alkylated counterpart would appear at m/z 360. Thus, we have good evidence that, in this particular peptide, the single alkylation site is uniquely Cys^3 (Cys^{11} in the original protein, see Figure 17). Thus, this peptide is tentatively rewritten as CVC*KEGGCK. The ill-resolved peaks at m/z 620-622 in Figure 19b correspond to the $\mathbf{b_5}$ and $\mathbf{y''_6}$ fragment ions predicted for this specifically alkylated peptide, while the peak at m/z 734 corresponds to the $\mathbf{b_7}$ ion. However, the peak at m/z 750 is not interpretable by any such simple scheme. It seems likely that this ion provides a further example of the kind of gross rearrangement observed recently [60] for the well-known peptide Substance P. This rearrangement involves wholesale transfer of residue(s) from the C-terminus to the ϖ-amino group of a Lys sidechain located somewhere near the N-terminus, with subsequent fragmentation of the rearranged peptide ion. In fact the fragment ion at m/z 750 corresponds closely to predictions based on a hypothesis that the $\mathbf{b_8}$ ion rearranges by transferring what is now the C-terminal Cys residue to the Lys sidechain, resulting in what is in effect a $\mathbf{b_7}$ ion in which the mid-chain Lys residue is derivatized by the Cys residue (additional 103 Da). The corresponding rearranged $\mathbf{a_6}$ fragment ion is predicted to appear at m/z 752.

The $(M + 2H)^{2+}$ ion (m/z 732) of the peptide of molecular mass 1460 Da, eluting at 11.5 min (Figures 13-15), was also fragmented in the same LC-MS/MS experiment.

The resulting fragment ion spectrum is shown in Figure 19c. This peptide was tentatively identified, in section E.2 above, as the tryptic double fragment (5,6) with all four Cys residues alkylated, *i.e.* as the dodecapeptide C*SPC*EKC*SSGC*K. However, the fragment ion spectrum shown in Figure 19c is difficult if not impossible to rationalize in terms of this postulated sequence. Almost none of the predicted **a**, **b**, and **y"** fragment ions are observed in the spectrum. However, the fragment ion sequence (1214 → 1054 → 966 → 806 → 646) is consistent with successive losses of Cys*, Ser, Cys* and Cys*. The only such sub-sequence present in the published sequence (Figure 17) is CSCC at residues 53-56, near the C-terminus. However, if these residues are to be lost **in the order observed** from a simple tryptic fragment of this protein (no disulfide bridges), they are being expelled from the N-terminus end towards the C-terminus, *i.e.* they must be **y"** ions. However, the remaining residues on the C-terminal side of the CSCC sub-sequence are only PT, with a combined molecular mass of 198 Da, so the combined mass of the [H-C*SC*C*PT-OH] portion is only 783 Da, much less than the 1214 Da from which the successive losses of C*, S, C* and C* are observed to occur (Figure 19c). Indeed, the only way in which an ion of molecular mass 1214 Da, derived from the published sequence and with no disulfide bridges, can lose these residues in the observed order is as the result of a large-scale rearrangement of the kind observed for Substance P [60], and described briefly above. Incidentally, the intense ions at m/z 248 and 408 are interpretable as (SC*) and (SC*)C*, respectively, where the order of the residues within parentheses is not determined.

It is possible that this rearrangement phenomenon is widespread in double (or higher) tryptic peptides containing in-chain (non-terminal) Lys residues. Certainly the single tryptic peptides (Lys or Arg at the C-terminus only) dissociated cleanly to complementary **y"** and **b** fragments, with only minor complications from doubly-charged fragments [58].

The peptide ion at m/z 671, eluting at 4.9 min (Figures 13-15), yielded the fragment ion spectrum shown in Figure 19a. This is a singly-charged precursor ion since no fragments were observed at m/z values above that of the precursor. (Any doubly-charged counterpart of the precursor, at m/z 336 would not have been observed since the LC-MS experiment (Figures 13-16) scanned down to only m/z 400). The intense ion at m/z 133 (Figure 21) again indicates presence of a Cys* residue. Note also that the pair of fragment ions at m/z 262 and 422 are separated by 160 Da, characteristic of a Cys* residue. It is difficult to deduce any further useful information from Figure 21.

The $(M + 2H)^{2+}$ ion from the peptide of molecular mass 2532 Da, eluting at 15.7 min, did not fragment significantly under the MS/MS conditions used. This is a large peptide in the present context of low-energy collision-induced dissociation, and this negative result is not unexpected.

In summary, only one of the tryptic fragments yielded an interpretable fragment ion

spectrum, and even here the interpretation was possible only in terms of consistency with an assumed sequence. That is, if this peptide had been a true unknown it would not have been possible to deduce its sequence from Figure 19b. Even so, this particular case was a success when compared with the results obtained for the other three peptides examined by MS/MS. These difficulties are all attributable to the fact that the Cys-alkylation and/or the tryptic digestion were incomplete, so that all four peptides were double (or possibly higher) tryptic peptides containing one or more highly basic Lys (or possibly Arg) residues within the chain (*i.e.* other than that at the peptide C-terminus), and possibly in corporating disulfide bridges. These features have several detrimental effects on the degree of sequence information obtained from MS/MS spectra of the $(M + 2H)^{2+}$ species. Localization of the protons on the basic sidechains prevents their facile mobilization along the peptide backbone for charge-site-induced fragmentation, and in the case of in-chain Lys residues such peptide ions are susceptible to wholesale rearrangements of residue(s) at the C-terminus to the Lys sidechain, thus making successful interpretation in terms of primary sequence much more difficult. Finally, formation of disulfide bridges prior to trypsin hydrolysis greatly complicates the interpretation.

F. SUMMARY AND CONCLUSIONS.

The work described in this report was intended as an exploratory study. The aim of this study was to determine whether techniques for protein characterization, based on modern methods of high resolution separation science and mass spectrometry, are applicable to metallothioneins in general and to those from marine species (*Scylla serrata*) in particular.

Metallothioneins are small proteins (6 kDa approximately) whose unique charateristics are a consequence of their very high cysteine content (approximately one-third of the 60 or so amino acid residues). This high Cys content accounts for the unique metal-binding capabilities of metallothioneins, and each molecule of the native protein is generally associated with 7 Cd^{2+} and/or Zn^{2+} ions. It was therefore not clear at the outset that techniques based on electrospray MS, CE-MS, LC-MS and LC-MS/MS of tryptic digests, would work well for characterization of metallothioneins.

The present exploratory experiments lead to a generally positive response to this question, but only if some precautions are taken. The present work has demonstrated, by HPLC and CE techniques, considerable heterogeneity in the vertebrate metallothionein isoforms prepared "pure" by conventional biochemical methods (though appreciably less heterogeneity in the case of *Scylla serrata*). In view of this feature, in addition to the potential for confusion arising from residual metal

ions, it is recommended that in any future work the metallothionein should first be demetallated, and the resulting apometallothionein fractionated by HPLC. Such pretreatment should greatly reduce the ambiguity of interpretation encountered in the present work. A suitable protocol for demetallation has been described in the literature [56].

Alkylation of the Cys residues, prior to specific proteolysis by trypsin or other enzymes, is an advisable step to avoid complications arising from disulfide bridge formation. The present work used iodoacetamide as the alkylating agent for the denatured but still metallated proteins. It is clear that this procedure was only partially successful, and that more vigorous reaction conditions would be required to fully alkylate the protein by this method. Hunziker [55] has persuasively argued the merits of cysteine modification by S-methylation using methyl-4-nitrobenzenesulfonate, and has described details of a procedure demonstrated to be successful for metallothioneins (S-methylation yields typically 98 ± 2%). While it is true that S-methylation is of marginal utility for purposes of counting Cys residues by comparing molecular masses before and after alkylation, such an experiment is of little use for metallothioneins in any case in view of the well-defined Cys requirements for formation of the metal-thiolate clusters.

Tryptic digestion of the partially alkylated metallothionein, as conducted in the present work, was disappointingly inefficient. Hunziker [49] has described experimental

conditions for successful digestion of S-methylated apometallothioneins by trypsin and other specific protoelytic enzymes. Fractionation of the alkylated apometallothioneins, prior to digestion, is recommended in order to minimize ambiguities of interpretation which inevitably arise when mixtures of closely related proteins are analyzed.

The disappointing results obtained here, for the MS/MS spectra of $(M + 2H)^{2+}$ ions from tryptic fragment peptides, are almost certainly the consequence of incomplete alkylation and/or tryptic digestion. All of the peptides thus studied were at least "double" tryptic fragments, *i.e.* contained at least one basic residue (mostly Lys) within the chain, in addition to that at the C-terminus. There are at least two consequences of this fact, both of them detrimental to the desired result of obtaining sequence information by MS/MS. The first of these is the preferred localization of the ionizing protons on the basic sidechains, rather than mobilization along the peptide backbone to initiate charge-site-induced fragmentations which yield sequence-specific fragment ions. The second harmful consequence of a mid-chain Lys residue involves the possiblility of large-scale rearrangement prior to fragmentation, which greatly complicates spectral interpretation. Of course, formation of disulfide bridges between non-alkylated Cys residues, in the alkaline medium which optimizes hydrolysis by trypsin, greatly increases the difficulty of deducing the primary sequence of an unknown. If the alkylation and digestion procedures can be carried through essentially to completion, as discussed above, it is anticipated that LC-MS/MS sequencing of the resulting tryptic peptides should be straightforward.

References:

[1] Y. Kojima, C. Berger, B.L. Vallee, J.H.R. Kägi, Proc. Natl. Acad. Sci. U.S.A. **73**, 3413 (1976)

[2] B.L. Vallee in "Methods of enzymology" vol. 205, J.F. Riordan and B.L. Vallee Eds., Acad. Press, NY, 1992, p.3

[3] D. Hamer, Annu. Rev. Biochem. **55**, 913 (1986)

[4] M. Margoshes, B.L. Vallee, J.Am. Chem. Soc. 79, 4813 (1957)

[5] M. Nordberg, Y. Kojima, Metallothionein Exp. Suppl. **34**, 41 (1979)

[6] J.H.R. Kägi, Y. Kojima, Metallothionein II Exp. Suppl. **52**, 25 (1987)

[7] W.E. Rauser, Annu. Rev. Biochem. **59**, 61 (1990)

[8] J.H.R. Kägi, B.L. Vallee, J. Biol. Chem. **235**, 3460 (1960)

[9] M.G. Cherian, R.A. Goyer, L. Delaquerrier-Richardson, Toxicol. Appl. Pharmacol., **38**, 399 (1976)

[10] K.S. Squibb, B.A. Fowler, Environ. Health Perspect. **54**, 31 (1984)

[11] M.G. Cherian in "Biological Roles of Metallothionein", E.C. Foulkes Ed., Elsevier, North-Holland, Amsterdam, 1982, p. 193

[12] C. Tohyama, N.Sugihira, H.Saito, J. Toxicol. Environ. Health, **22**, 255 (1987)

[13] J.H.R. Kägi in "Methods of Enzymology" vol. 205, J.F. Riordan and B.L. Vallee Eds., Acad. Press, NY, 1992, p. 613

[14] M. Webb in Metallothionein II, Experientia Suppl. **52**, 483 (1987)

[15] J.H.R. Kägi, A. Schäffer, Biochemistry, **27**, 8509 (1988)

[16] B.L. Vallee, D.S. Auld, Biochemistry, **29**, 5647 (1990)

[17] F.O. Brandy, Trends Biochem. Sci. **7**, 143 (1982)

[18] M. Beltramini, K. Lerch, FEBS Lett., **142**, 219 (1982)

[19] B.L. Vallee, Metallothionein Exp. Suppl. **34**, 19 (1979)

[20] K. Marin, Cell, **41**, 9 (1985)

[21] J.H.R. Kägi, M. Vasák, K. Lerch, D.E.O. Gilg, P. Hunziker, W.R. Bernhard, M. Good, Environ. Health Perspect., **54**, 93 (1984)

[22] J.H.R. Kägi, Y. Kojima, in Metallothionein II, Experientia Suppl., **52**, 25 (1987)

[23] D.R. Winge, K.B. Nielson, W.R. Gray, D.H. Hamer, J. Biol. Chem. **260**, 14464 (1985)

[24] Y. Hayashi, C.W. Nakagawa, A. Murasugi, Environ. Health Perspect. **65**, 13 (1986)

[25] W.R. Bernhard, J.H.R. Kägi, in Metallothionein II, Experientia Suppl. **52**, 309 (1987)

[26] E. Grill, E.L. Winnacker, M.H. Zenk, Proc. Natl. Acad. Sci. U.S.A., **84**, 439 (1987)

[27] W.E. Rauser, Annu. Rev. Biochem. **59**, 61 (1990)

[28] N.J. Robinson, P.J. Jackson, Physiol. Plant, **67**, 499 (1986)

[29] K. Lerch, D. Ammer and R.W. Olafson, J. Biol. Chem. **257**, 2420 (1982)

[30] J.D. Otvos, R.W. Olafson and I.M. Armitage, J. Biol. Chem.**257**, 2427 (1982).

[31] M.Barber, R.S. Bordoli, R.D. Sedgwick and A.N. Tyler, J. Chem. Soc. Chem. Commun. 324 (1981).

[32] M. Barber and B.N. Green, Rapid Commun. Mass Spectrom. **1**, 80 (1987).

[33] K. Biemann and S.A. Martin, Mass Spectrom. Rev. **6**, 1 (1987).

[34] R.M. Caprioli, T. Fan and J.S. Cottrell, Anal. Chem. **58**, 2949 (1986).

[35] R.D. Smith, J.A. Loo, C.G. Edmonds, C.J. Barinaga and H.R. Udseth, Anal. Chem. **62**, 882 (1990).

[36] A.P. Bruins, T.R. Covey and J.D. Henion, Anal. Chem. **59**, 2642 (1987).

[37] P. Thibault, C. Paris and S. Pleasance, Rapid Commun. Mass Spectrom. **5**, 484 (1991).

[38] "Metallobiochemistry, Part B: Metallothionein and Related Molecules", Methods in Enzymology, Vol. 205, (J.F. Riordan and B.L. Vallee, Eds.), Academic Press, New York, 1991, Chaps. 9-21.

[39] C.D. Klaassen and L.D. Lehman-McKeeman, in "Metallobiochemistry, Part B: Metallothionein and Related Molecules", Methods in Enzymology, Vol. 205, (J.F. Riordan and B.L. Vallee, Eds.), Academic Press, New York, 1991, Chap. 22

[40] K.T. Suzuki, in "Metallobiochemistry, Part B: Metallothionein and Related Molecules", Methods in Enzymology, Vol. 205, (J.F. Riordan and B.L. Vallee, Eds.), Academic Press, New York, 1991, Chap. 23.

[41] M.P. Richards, in "Metallobiochemistry, Part B: Metallothionein and Related Molecules", Methods in Enzymology, Vol. 205, (J.F. Riordan and B.L. Vallee, Eds.), Academic Press, New York, 1991, Chap. 25

[42] R.W. Olafson and P.E. Olsson, in "Metallobiochemistry, Part B: Metallothionein and Related Molecules", Methods in Enzymology, Vol. 205, (J.F. Riordan and B.L. Vallee, Eds.), Academic Press, New York, 1991, Chap. 24.

[43] J.H. Beattie, M.P. Richards and R. Self, J. Chromatogr. **632**, 127 (1993).

[44] J.W. Jorgenson, Anal. Chem. **58**, 743A (1986).

[45] A.G. Ewing, R.A. Wallingford and T.M. Olefirowicz, Anal. Chem. **61**, 292A (1989).

[46] P.D. Grossman *et al.*, American Biotech. Lab., Feb. 1990, pp. 35-43

[47] S. Pleasance, P. Thibault and J. Thompson, Proc. 38th Annual Conf., Amer. Soc. Mass Spectrom., (Tucson, 1990), pp. 720-721.

[48] X. Yu, M. Wojciechowski and C. Fenselau, Anal. Chem. **65**, 1355-1359 (1993).

[49] P.E. Hunziker, in "Metallobiochemistry, Part B: Metallothionein and Related Molecules", Methods in Enzymology, Vol. 205, (J.F. Riordan and B.L. Vallee, Eds.), Academic Press, New York, 1991, Chap. 49.

[50] J.A. Olivares, H.T. Nguyen, C.R. Yonker and R.D. Smith, Anal. Chem. **59**, 1230 (1987).

[51] R.D. Smith, J.A. Olivares, H. Nguyen and H.R. Udseth, Anal. Chem. **60**, 436 (1988).

[52] S. Pleasance, P. Thibault and J.F. Kelly, J. Chromatogr. **591**, 325 (1992).

[53] J. Zaia, R.S. Annan and K. Biemann, Rapid Commun. Mass Spectrom. **6**, 32 (1992).

[54] P. Thibault, S. Pleasance, M.V. Laycock, R.M. Mackay and R.K. Boyd, Int. J. Mass Spectrom, Ion Proc. **111**, 317 (1991).

[55] P.E. Hunziker, in "Metallobiochemistry, Part B: Metallothionein and Related Molecules", Methods in Enzymology, Vol. 205, (J.F. Riordan and B.L. Vallee, Eds.), Academic Press, New York, 1991, Chap. 45.

[56] M.M. Vestling, C.M. Murphy and C. Fenselau, Anal. Chem. **62**, 2391 (1990).

[57] P. Roepstorff and J. Fohlman, Biomed. Env. Mass Spectrom. **11**, 601 (1984).

[58] X.-J. Tang and R.K. Boyd, Rapid Commun. Mass Spectrom. **6**, 651 (1992).

[59] D. Zidarov, P. Thibault, M.J. Evans and M.J. Bertrand, Biomed. Env. Mass Spectrom. **19**, 13 (1990).

[60] X.-J. Tang, P. Thibault and R.K. Boyd, Anal. Chem. **65**, 2824 (1993).

Table 1. Amino acid sequences for metallothionein isoforms with calculated and experimental molecular weights. (Underlined sequences are uncertain, Ac indicates an acetylated N-terminus, X indicates an undetermined blocked N-terminus). Adapted from Ref. 6, which gives original reference.

	Metallothionein	Amino acid sequence	No. of Basic Residues[a]	Calculated Mol. Wt.[b]	Observed Mol. Wt.[c]	Observed Max charge state
		1 10 20 30 40 50 60				
		l l l l l l l				
Vertebrate	Human MT-I	MDP.NCSCAT.GGSCTC.TGSCKC.KECKCNSCKKSCC.SCCPMSC.....AK..CAQGCICKG.ASE.KCSCCA	8	6133.4		
	Human MT-II	MDP.NCSCAA.GDSCTC.AGSCKC.KECKCTSCKKSCC.SCCPVGC.....AK..CAQGCICKG.ASD.KCSCCA	8	6042.3		
	Monkey MT-I	MDP.NCSCAT.GVSCTC.ADSCKC.KECKCTSCKKSCC.SCCPVGC.....AK..CAQGCVCKG.ASD.KCNCCA	8	6127.4		
	Monkey MT-II	MDP.NCSCVA.GDSCTC.AGSCKC.KECKCTSCKKSCC.SCCPVGC.....AK..CAQGCICKG.ASD.KCNCCA	7	6097.4		
	Horse MT-IA	Ac-MDP.NCSCPT.GGSCTC.AGSCKC.KECRCTSCKKSCC.SCCPGGC.....AR..CAQGCVCKG.ASD.KCSCCA (R; L)	7	6181.4 / 6108.4 / 6082.3	6081.3	+7
	Horse MT-IB	Ac-MDP.NCSCVA.GESCTC.AGSCKC.KQCRCASCKKSCC.SCCPVGC.....AK..CAQGCVCKG.ASD.KCSCCA	7	6109.4	6110.6	+7
	Bovine MT-II	MDP.NCSCPT.GGSCSC.AGSCTC.KACRCPSCKKSCC.SCCPVGC.....AK..CAQGCVCKG.ASD.KCSCCA	7	5951.2		
	Rabbit MT-I	Ac-MDP.NCSCAA.BGSCTC.ATSCRC.KECKCTSCKKSCC.SCCPAGC.....TK..CAQGCICKG.ASD.KCSCCA (L)	7	6069.4 / 6043.3		+7
	Rabbit MT-II	Ac-MDP.NCSCAA.DGSCTC.ATSCKC.KECKCTSCKKSCC.SCCPSGC.....AK..CAQGCICKG.ASD.KCSCCA (A)	8	6058.3 / 6116.3		+7
	Rabbit MT-IIA	X-MDP.NCSCAAAGDSCTC.ANSCTC.KACKCTSCKKSCC.SCCPPGC.....AK..CAQGCICKG.ASD.KCSCCA	7	6083.3		+7
	Rabbit MT-IIB	X-MDP.NCSCAT.GDSCTC.ASSCKC.KECKCTSCKKSCC.SCCPAGC.....TK..CAQGCICKG.ASD.KCSCCA	8	6104.3		+7
	Chicken MT	MDPQDCTCAA.GDSCSC.AGSCKC.KNCRCRSCRKSCC.SCCPAGC.....NN..CAKGCVCKEPASS.KCSCCH	10	6461.7		
	Trout MT-I	MDP..CECSK.TGSCNC.GGSCKC.SNCACTSCKKSCC.DCCPSGC.....SK..CASGCVCKGKTCD.T.SCCQ	7	5993.1		
	Trout MT-II	MDP..CECSK.TGSCNC.GGSCKC.SNCACTSCKKSCC.PCCPSDC.....SK..CASGCVCKGKTCD.T.SCCQ	7	6033.2		
Invertebrate	Crab MT-I	PGP..C.C...NDKCVCKEGGCK..EGCQCTSCRCSPC.EKCSSGC.....K..CANKEECSKTCSK.ACSCCPT	9	6001.1	6057.6	+7
	Crab MT-II	PDP..C.C...NDKCDCKEGECK..TGCKCTSCRCPPC.EQCSSGC.....K..CANKEDCRKTCSK.PCSCCP	9	6109.2		
	Oyster MT	MV.CKCDCKNQNCSCNTGTKDCDC.SDAKC..CEQYCCPTASEKKC.....CKSGCAGGCKCAN.CEC.AQAAH	5	6593.7		
	Sea Urchin MT	MPDVKCVCCTEGKECACFGQDC.C.VTGEC..CKDGTCCGICTNAA.....CK..CANGCKCGSGCSC.TEGNCAC	5	6388.5		
	Fruitfly MT	MPCPC.GSGCKCASQTK.GSCN.....CGSDCKCGGDKKS.ACGCSE	5	3782.4		
	Yeast MT	MFSELINFQN.EGHCQCQCGSCKN.NE....QCQKSCS...CPTGC.....NS...DDKCPCGNKSEETKKSCCSGK	7	6521.3		
Plant	Soybean MT	MSCCGGNCGC.GSSCKC.GNGCGGCKMYPDLSYTESTTTETLVMGVAPVKAQFQGAEMGVPAENDGCKCGPNCSCNPCTCK	4	7938.2		
	Pea MT	MS....GCGC.GSSCNC.GDSCKCNKRSSGLSYSQMQTTQTVILGVGPAKIQFQGAQMSAASQGDGCLCGDNCTCDPCNCK	4	7587.7		
	Wheatgerm MT	GCNDKCGC.AVPCPG.GTGCRCTSANSGAAAGEHTTCGC.....GEHCGGNPCACGGEGTPSGCAN	2	5380.9		

a lysine, arginine and histidine residues

b average molecular weight

c considering both basic amino acids and terminal blocking groups

d calculated mol weight independent of unknown terminus

Table 2. Comparison of molecular weights (Da) deduced from electrospray mass spectra, for rabbit liver metallothionein isoforms I and II, as a function of pH.

Isoform I		Isoform II	
pH=2.8	**pH=4.5**	**pH=2.8**	**pH=4.5**
6144	6609	6125	6593
6174	6654	6146	6655
6214	6684	6154	6683
6241	6719	6187	6714
	6880	6215	
		6241	
		6269	
		6355	

CAPTIONS FOR FIGURES

Figure 1. Chromatograms of metallothionein preparations obtained by HPLC/UV (214 nm). Chromatographic conditions were those described in the Experimental Methods section.

(a) Rabbit liver metallothionein isoform I.

(b) Rabbit liver metallothionein isoform II.

(c) Crab metallothionein isoform I.

(d) Mixtures of isoforms of horse metallothionein.

Figure 2. Diagram showing the mechanism of electroendosmosis, indicating the negative charge on the fused silica capillary wall, the positively charged double layer adjacent to the wall, the electroendosmotic velocity V_{eo}, and the electrophoretic velocity V_{eo}. At high pH the negatively charged analyte species undergo electrophoretic migration towards the positive electrode, but at sufficiently low pH they acquire a positive charge and the electrophoretic velocity is directed towards the negative electrode. The direction of the electroendosmotic flow is always toward the negative electrode, for untreated silica capillaries, but the net concentration of positive charge in the double layer, and thus V_{eo}, varies with pH.

Figure 3. Block diagram of apparatus for capillary electrophoresis.

Figure 4. Electropherograms, obtained using CE with UV detection at 214 nm, of rabbit liver metallothionein preparations. Sodium citrate buffer, pH 2.0; 30 kV potential difference across the capillary.

(a) unseparated isoforms; (b) isoform I; (c) isoform II.

Figure 5. Electrospray mass spectrum of rabbit liver metallothionein isoform I in aqueous formic acid at pH 2.8, introduced by flow injection.

(a) Complete spectrum.

(b) Expanded region, m/z 1010-1060.

(c) Complete spectrum transformed from m/z scale to molecular weight scale.

Figure 6. Electrospray mass spectrum of rabbit liver metallothionein isoform II in aqueous formic acid at pH 2.8, introduced by flow injection.

(a) Complete spectrum.

(b) Expanded region, m/z 1010-1060.

(c) Complete spectrum transformed from m/z scale to molecular weight scale.

Figure 7. Electrospray mass spectra of three metallothionein preparations, each in aqueous formic acid at pH 2.8, introduced by flow injection.

(a) Crab (*Scylla serrata*).

(b) Equine kidney.

(c) Rabbit liver isoform II.

Figure 8. LC-MS analysis (SIM mode) of a commercial preparation of rabbit liver metallothionein isoform I, used as supplied. Experimental conditions described in text.

(a) LC-UV chromatogram (same as Figure 1a).

(b) LC-MS chromatogram, sum of intensities of eight monitored ions (m/z 878, 883, 888, 892, 1025, 1030, 1036 and 1041).

(c) LC-MS chromatogram for m/z 878.

(d) LC-MS chromatogram for m/z 1025.

Figure 9. LC-MS analysis (SIM mode) of a commercial preparation of rabbit liver metallothionein isoform I, used as supplied. Experimental conditions described in text. The individual m/z values monitored are noted on each chromatogram. Those for m/z 878 and 1025 are the same as those shown in Figures 8c and 8d, respectively. The relative intensities of the six different chromatograms are meaningful. The small peaks in the chromatograms for m/z 888 and 1036 represent interferences from the isotopic ion clusters centred at m/z 883 and 1030 (not shown).

Figure 10. Schematic diagram of the articulated IonSpray® interface for the SCIEX API III mass spectrometer, configured as a CE-MS interface. The solid black line at the top left represents the CE capillary, which terminates as shown at the nebulizer needle.

Figure 11. Analysis by CE-MS of isoform I of metallothionein from crab (*Scylla serrata*). Approximately 16 ng of protein were pressure-injected. The CE buffer was acetic acid, pH 2.2. The mass spectrometer was scanned from m/z 750 to 1400, with a step size of 1 m/z unit and a dwell time per step of 5 ms.

(a) Total ion electropherogram.

(b) Mass spectrum obtained at crest of CE peak at 14.7 min. The insert is the spectrum converted to the molecular weight scale.

Figure 12. LC-MS analysis of the reaction product mixture from denaturation of crab metallothionein isoform I, with subsequent reduction and alkylation of the Cys thiol groups. The LC conditions used were the same as those used to obtain Figures 8 and 9, but the mass spectrometer was scanned over the range m/z 800 to 1600.

Figure 13. Total ion chromatogram (m/z 400 to 1400) obtained by LC-MS analysis of the tryptic digest of the alkylated component eluting at 23.0 min in Figure 12 (of molecular weight 6580 Da). The HPLC gradient was 5 to 50% acetonitrile in water in 30 min, with 0.1% TFA.

Figure 14. Two-dimensional contour representation (ion intensity represented by depth of shading, as a function of both m/z and of retention time) of the same LC-MS analysis of the tryptic digest of the alkylated product of crab metallothionein which eluted at 23.0 min in Figure 12. This is an alternative representation of the same experiment as that summarized in Figure 13.

Figure 15. Electrospray mass spectra obtained for LC peaks shown in Figures 13 and 14. The retention time (min) at which a spectrum was obtained is shown at the top left of that spectrum (6.4, 12.5, 11.5 and 4.9 min, reading from top to bottom).

Figure 16. Electrospray mass spectra obtained for LC peaks shown in Figures 13 and 14. The retention time (min) at which a spectrum was obtained is shown at the top left of that spectrum (13.3, 15.7 and 16.3 min, reading from top to bottom).

Figure 17. Molecular masses predicted for tryptic fragment peptides of crab metallothionein isoform I, according to the published sequence [6,29]. Output from the software package "MacProMass".

Figure 18. Molecular masses predicted for tryptic fragment peptides of crab metallothionein isoform I, according to the published sequence [6,29] but assuming all Cys residues are alkylated ($-CH_2$ - C0 - NH_2 replacing -H). Output from the software package "MacProMass".

Figure 19. Fragment ion spectra, obtained by an LC-MS/MS, experiment, of peptide ion precursors identified previously by LC-MS (Figures 13-16). Mass spectrometric conditions are described in the Experimental section. The precursor ions and their elution times are:

(a) m/z 671 (4.9 min);

(b) m/z 492 (6.4 min);

(c) m/z 732 (11.5 min).

Figure 1

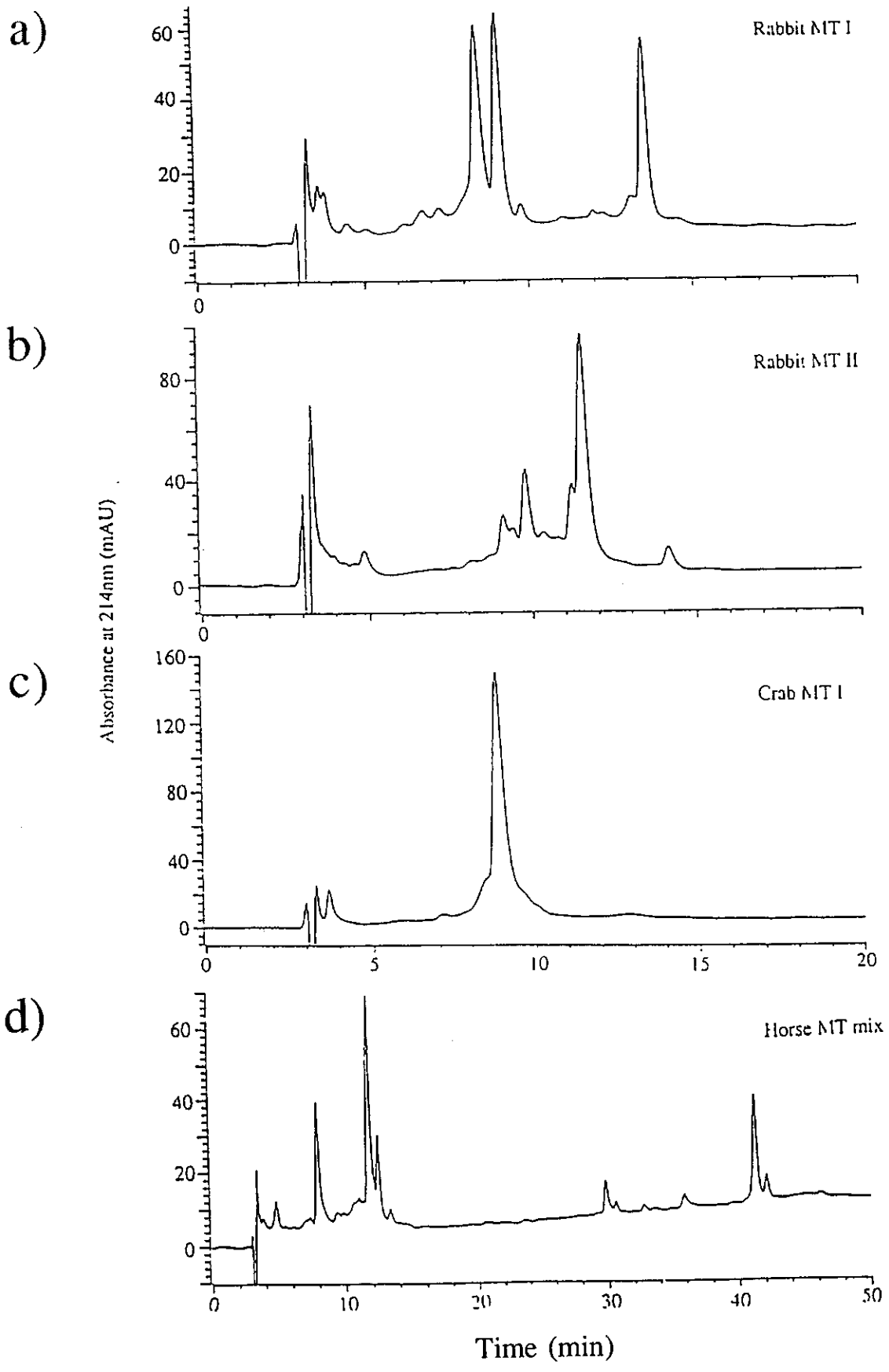

a)
Rabbit MT I
b)
Rabbit MT II
c)
Crab MT I
d)
Horse MT mix
Absorbance at 214nm (mAU)
Time (min)

Figure 2

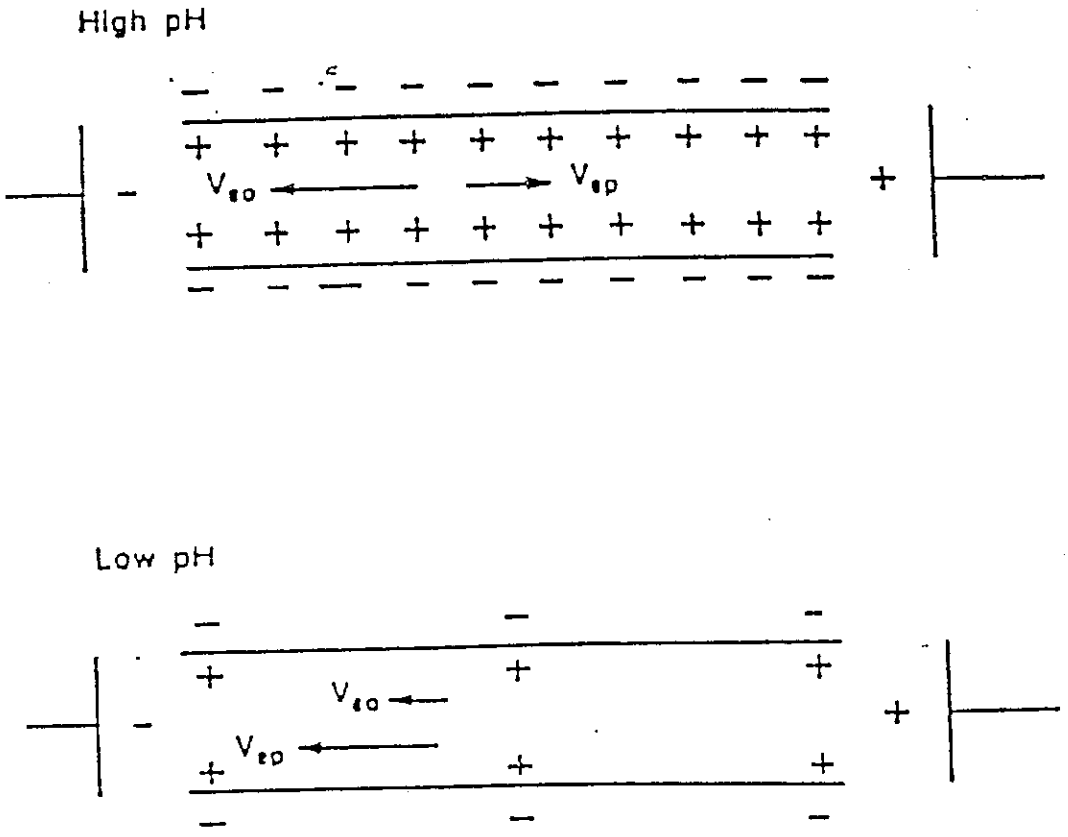


Figure 3

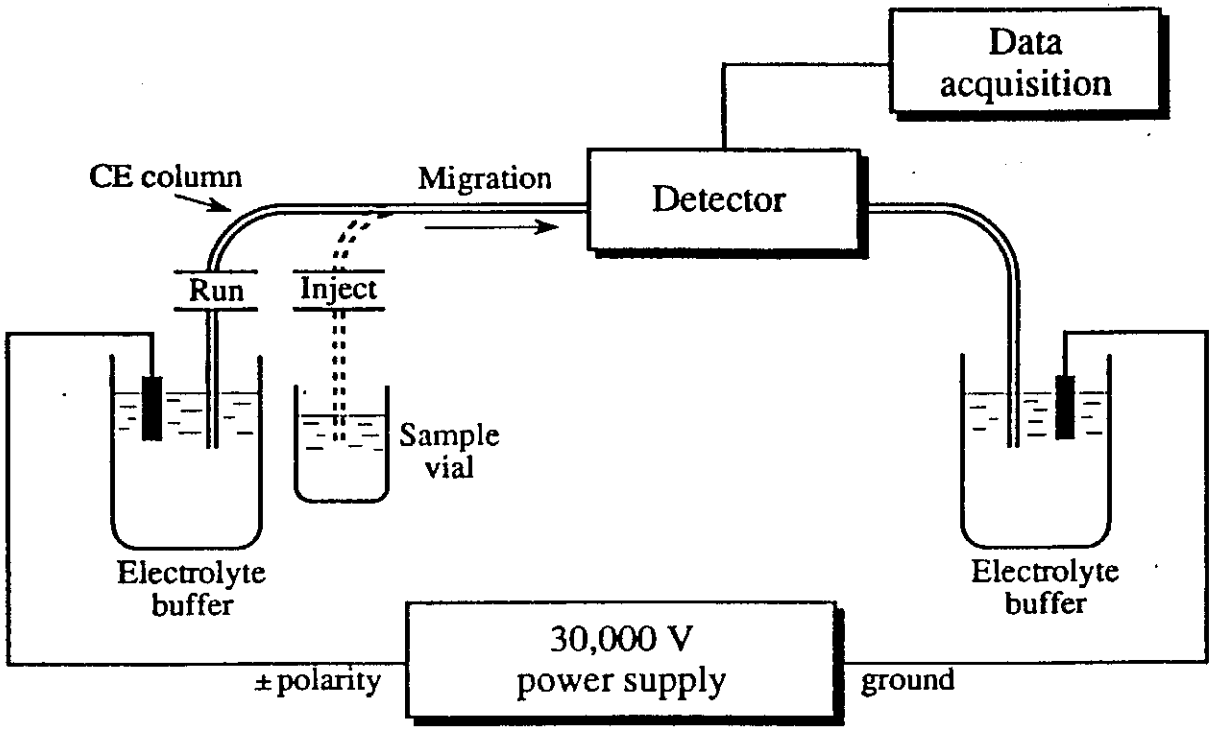
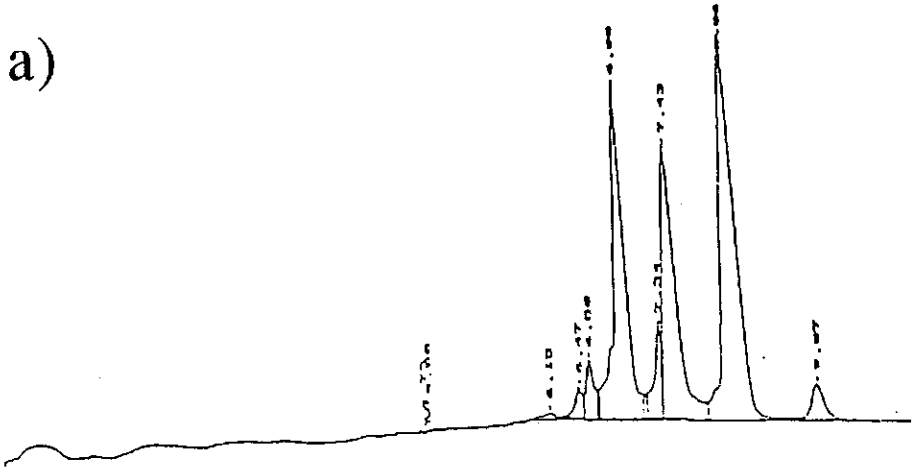
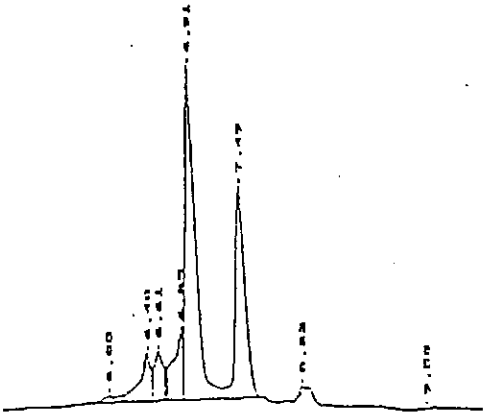

Data acquisition
CE column
Migration
Detector
Run
Inject
Sample vial
Electrolyte buffer
Electrolyte buffer
30,000 V power supply
± polarity
ground

Figure 4

a)



b)



c)

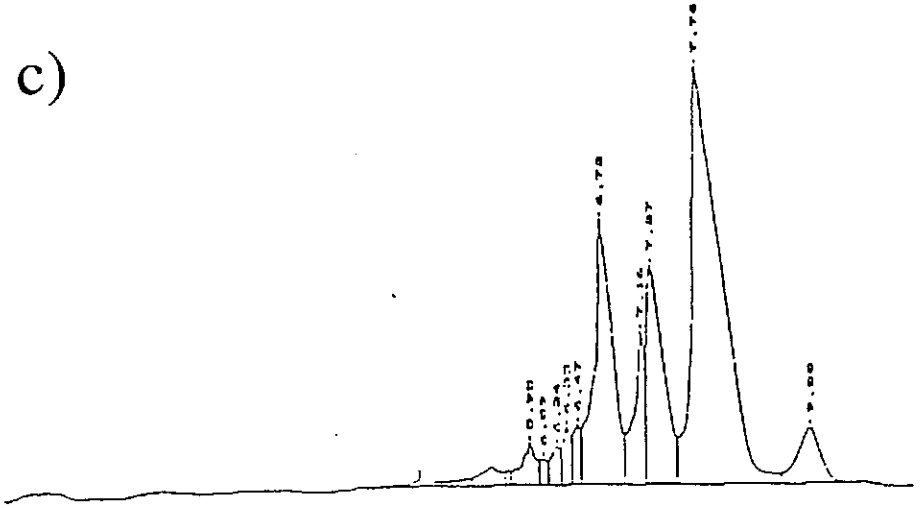
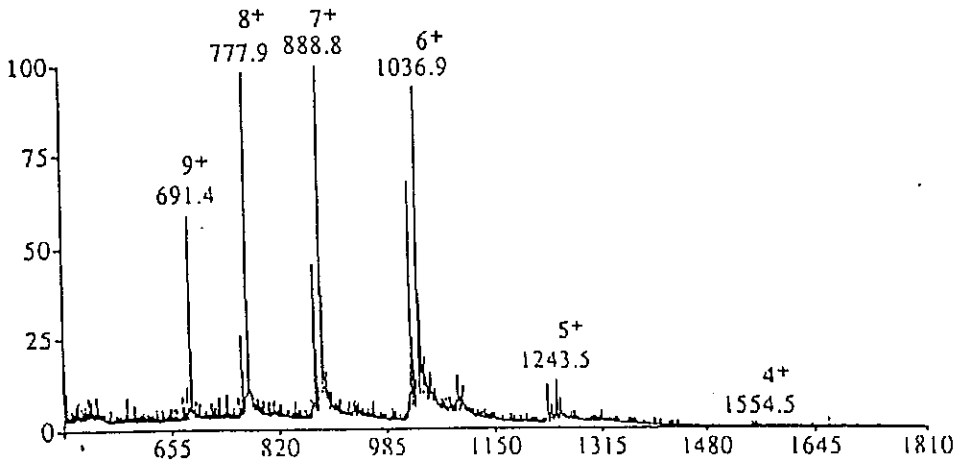
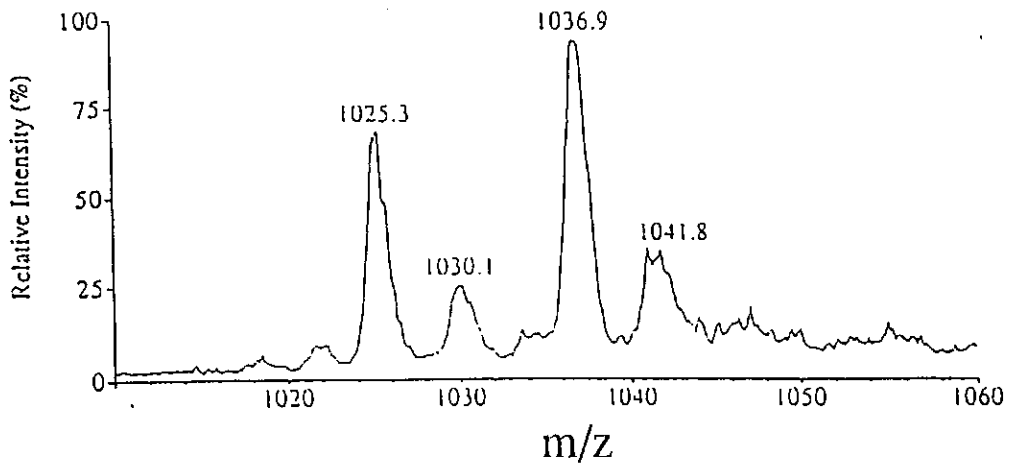


Figure 5

a)

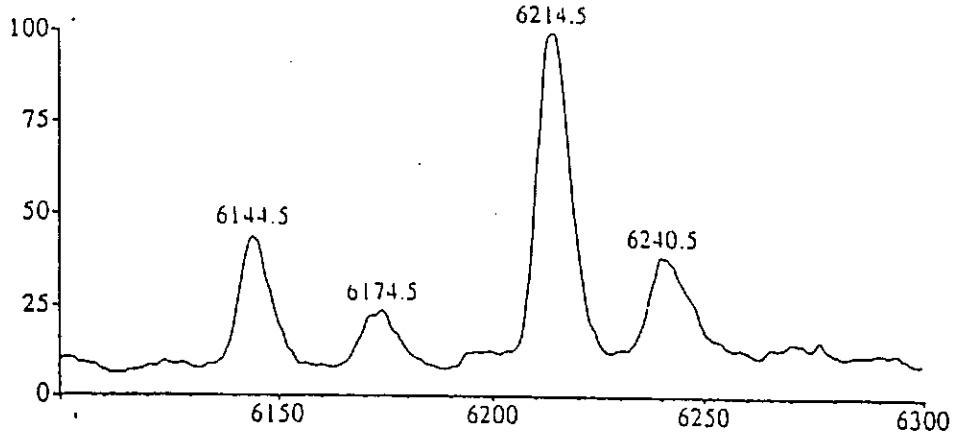


b)



m/z

c)



Mol. Wt.

Figure 6

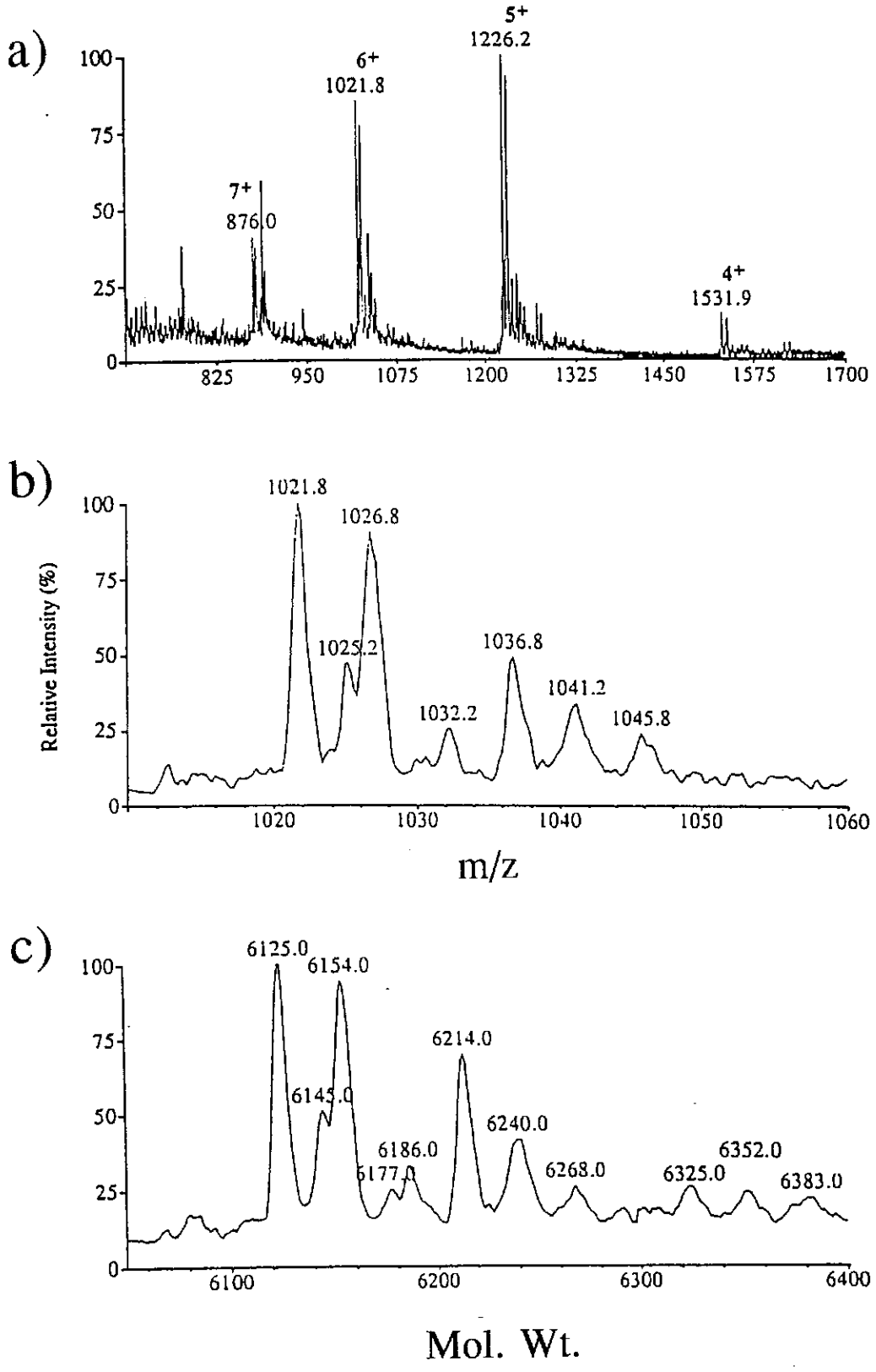

a)
100
75
50
25
0
7+
876.0
6+
1021.8
5+
1226.2
4+
1531.9
825
950
1075
1200
1325
1450
1575
1700
b)
Relative Intensity (%)
1021.8
1026.8
1025.2
1032.2
1036.8
1041.2
1045.8
1020
1030
1040
1050
1060
m/z
c)
6125.0
6154.0
6214.0
6145.0
6240.0
6186.0
6177.0
6268.0
6352.0
6325.0
6383.0
6100
6200
6300
6400
Mol. Wt.

Figure 7

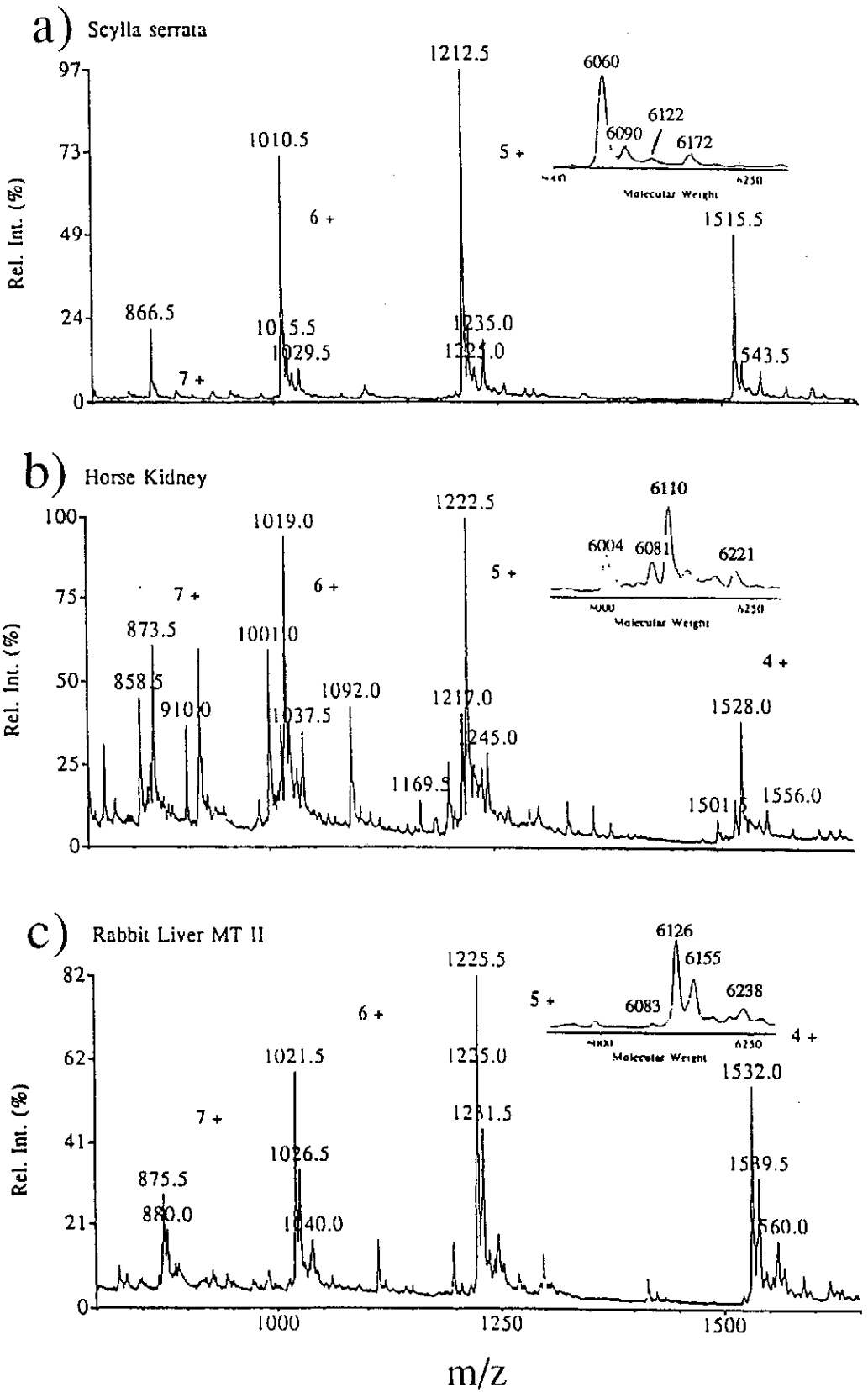

a) Scylla serrata
Rel. Int. (%)
97
73
49
24
0
1212.5
1010.5
6 +
5 +
6060
6122
6090
6172
Molecular Weight
1515.5
866.5
7 +
1015.5
1029.5
1235.0
1221.0
543.5
b) Horse Kidney
Rel. Int. (%)
100
75
50
25
0
1222.5
1019.0
6110
6004
6081
6221
Molecular Weight
7 +
6 +
5 +
4 +
873.5
858.5
910.0
1001.0
1092.0
1037.5
1217.0
1245.0
1528.0
1169.5
1556.0
c) Rabbit Liver MT II
Rel. Int. (%)
82
62
41
21
0
1225.5
6126
6155
6238
6083
5 +
4 +
Molecular Weight
6 +
1021.5
1235.0
1532.0
7 +
1231.5
1026.5
875.5
880.0
1040.0
1589.5
1560.0
1000
1250
1500
m/z

Figure 8

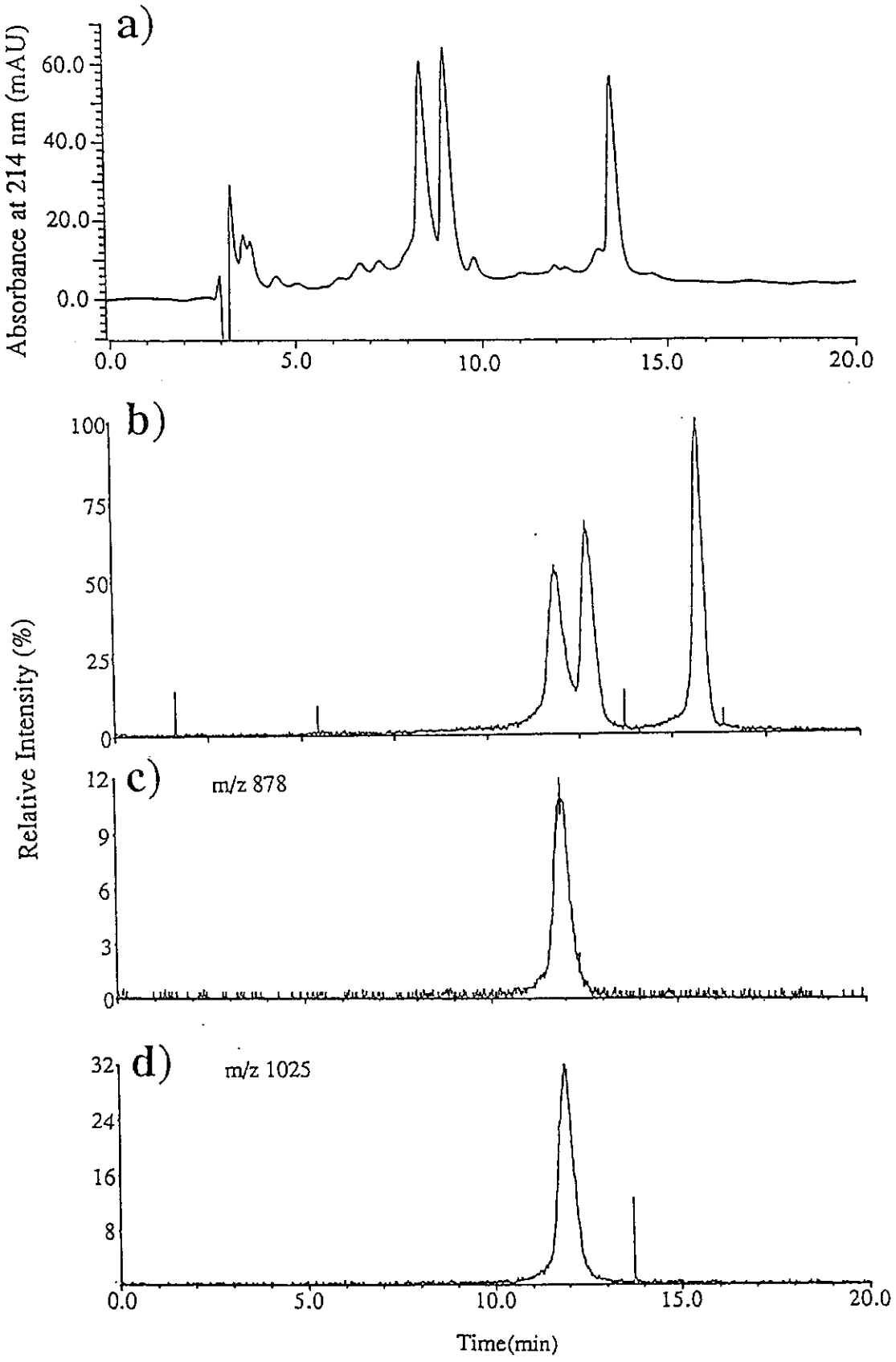
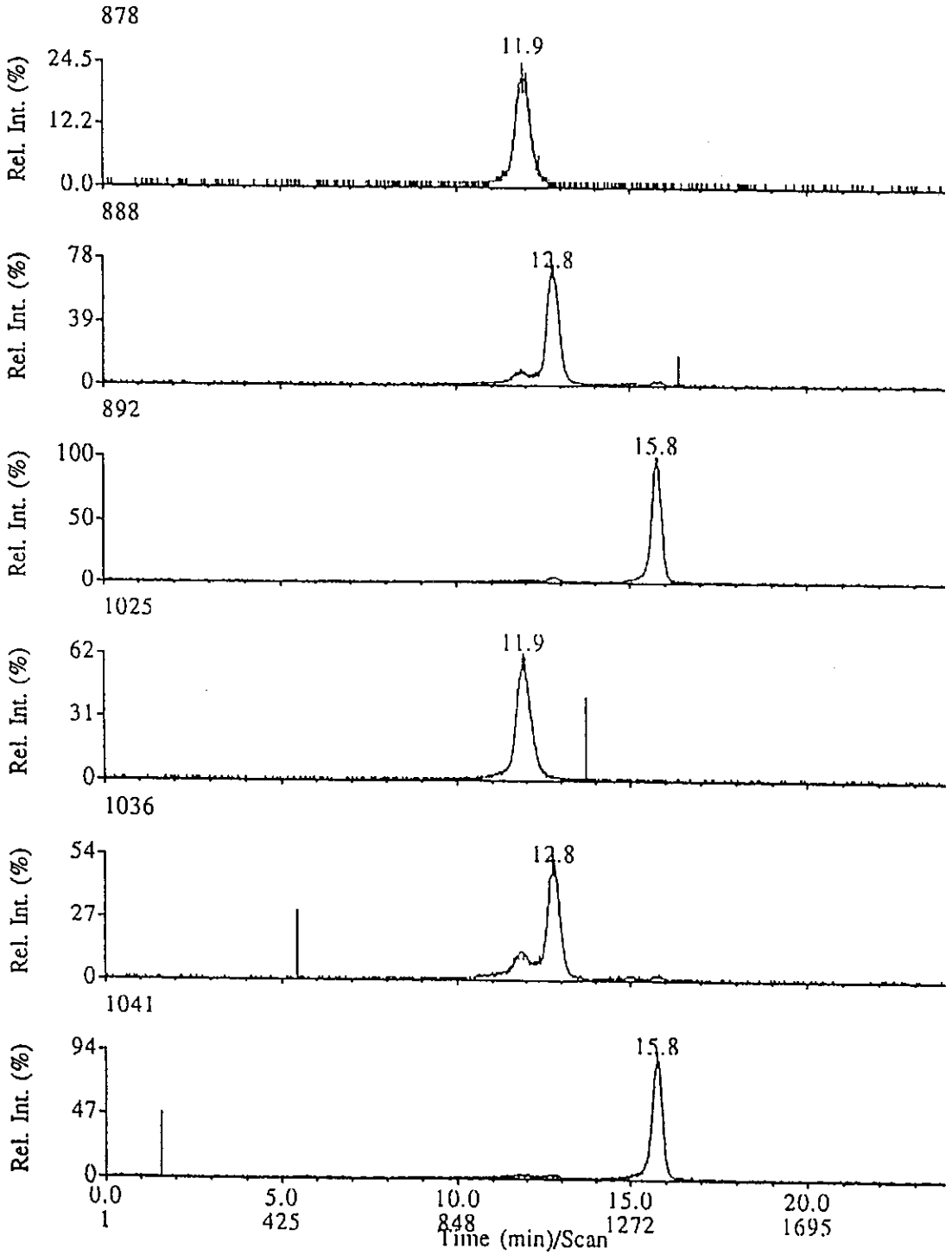

a)
Absorbance at 214 nm (mAU)
60.0
40.0
20.0
0.0
0.0
5.0
10.0
15.0
20.0
b)
100
75
50
25
0
c)
m/z 878
12
9
6
3
0
d)
m/z 1025
32
24
16
8
0.0
5.0
10.0
15.0
20.0
Relative Intensity (%)
Time(min)

Figure 9


878
11.9
24.5
12.2
0.0
Rel. Int. (%)
888
78
39
0
12.8
892
100
50
15.8
1025
62
31
11.9
1036
54
27
12.8
1041
94
47
15.8
0.0
5.0
10.0
15.0
20.0
1
425
848
1272
1695
Time (min)/Scan

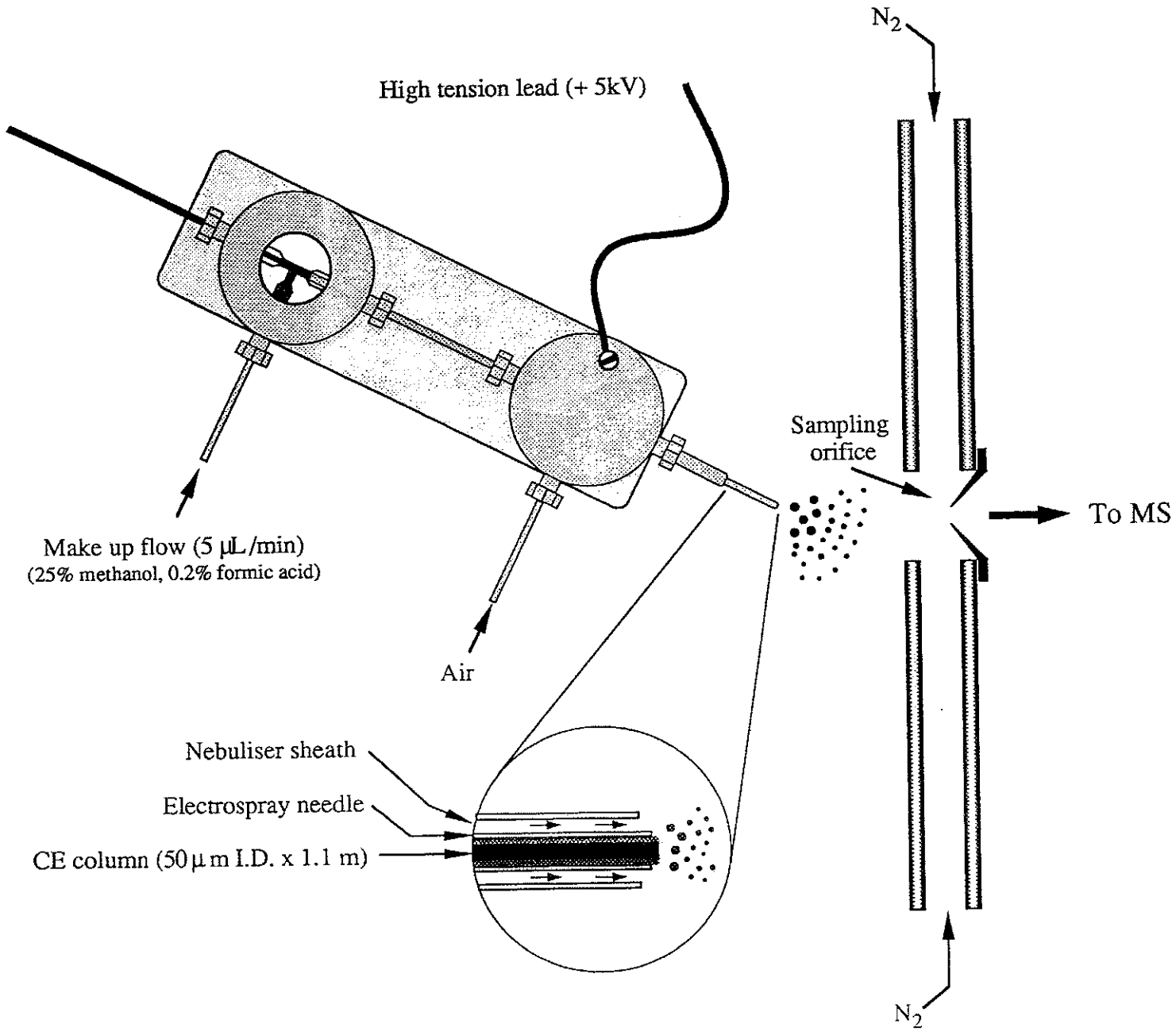
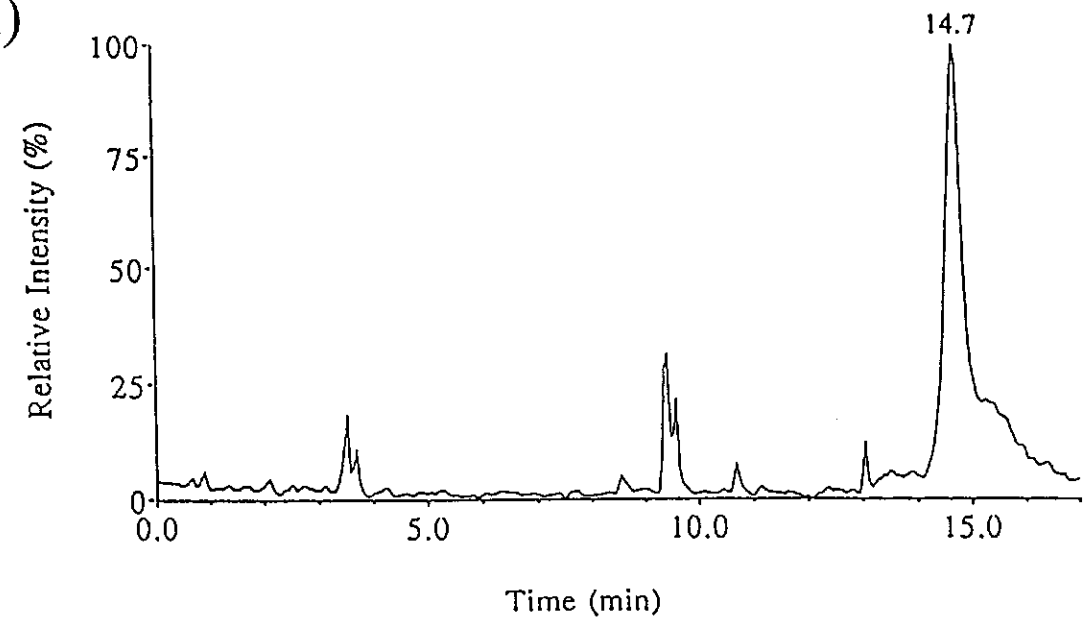

N_2
High tension lead (+ 5kV)
Sampling orifice
To MS
Make up flow (5 µL/min)
(25% methanol, 0.2% formic acid)
Air
Nebuliser sheath
Electrospray needle
CE column (50 µm I.D. x 1.1 m)
N_2

Figure 10

Figure 11

a)



Full mass spectrum of peak at 14.7 min.

b)

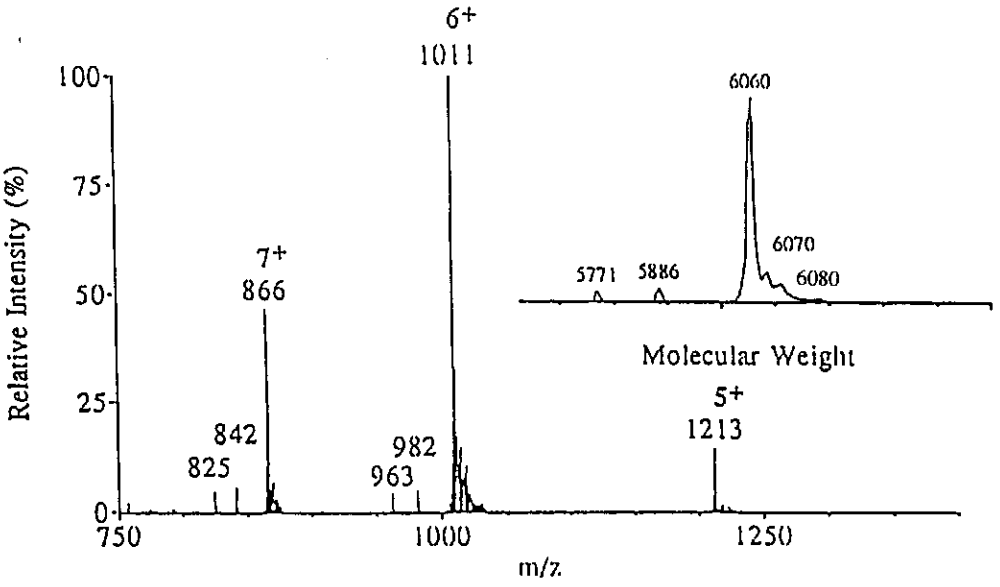


Figure 12

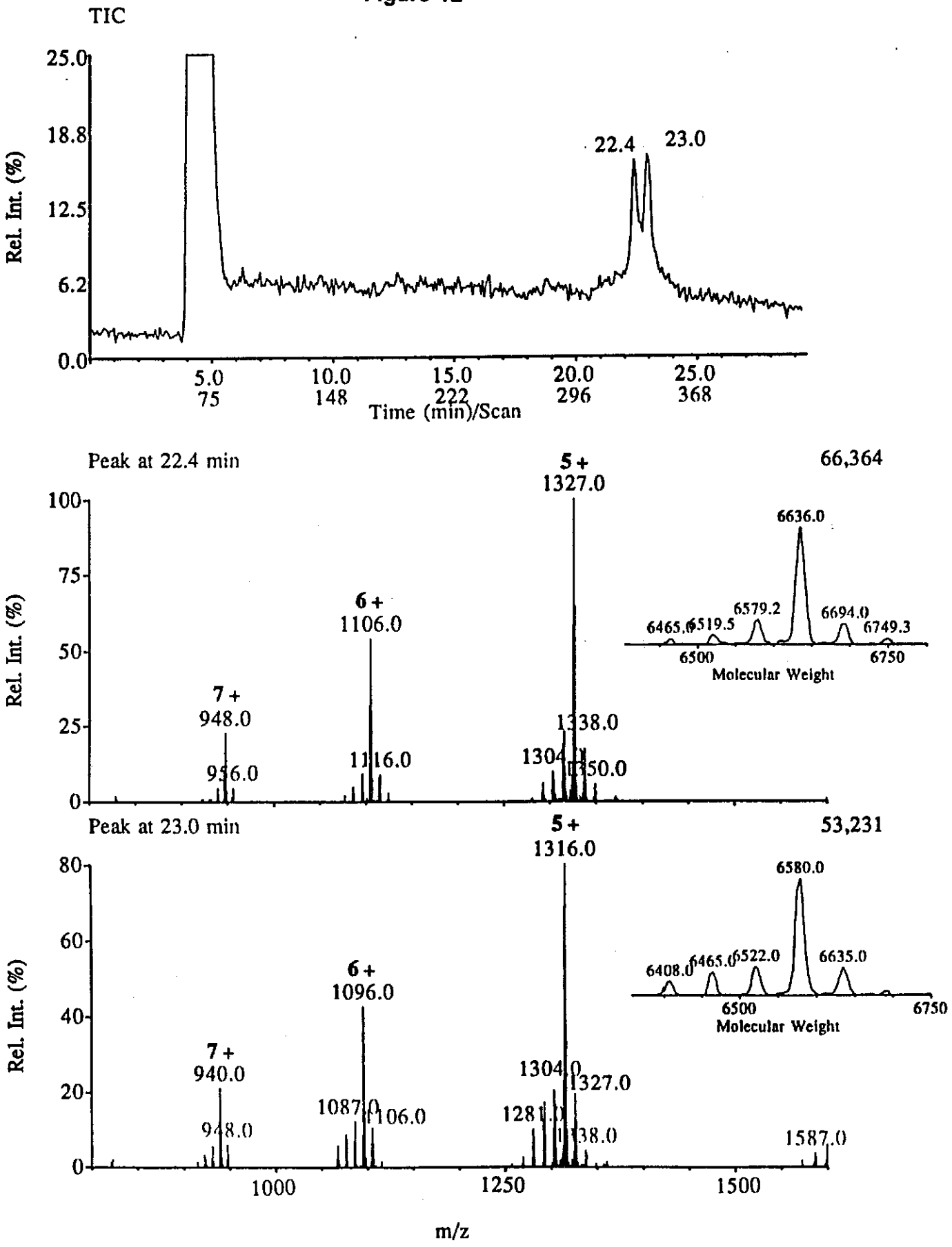
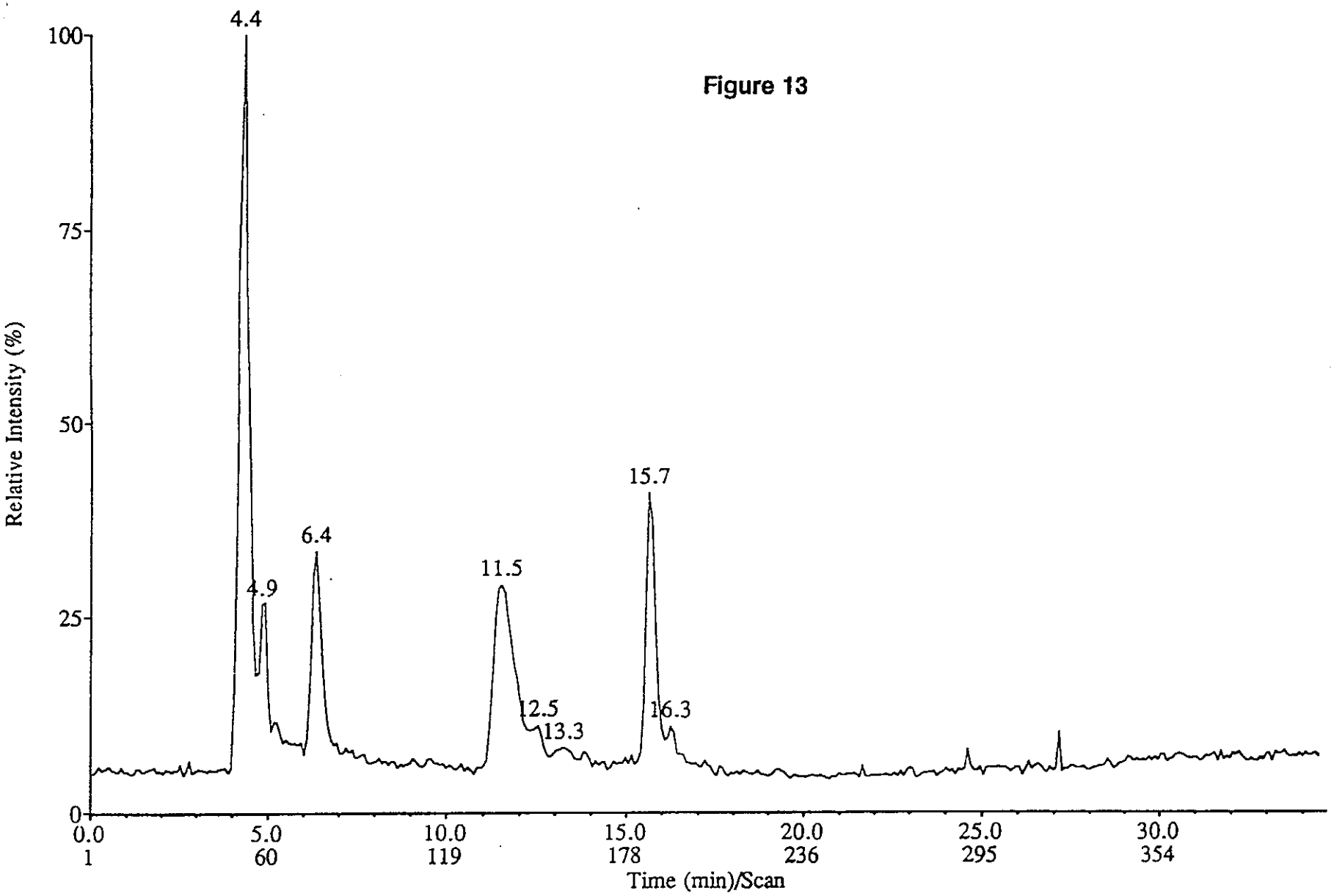

TIC
25.0
18.8
12.5
6.2
0.0
Rel. Int. (%)
22.4
23.0
5.0
75
10.0
148
15.0
222
20.0
296
25.0
368
Time (min)/Scan
Peak at 22.4 min
66,364
5+
1327.0
6+
1106.0
7+
948.0
956.0
1116.0
1304.0
1338.0
1350.0
6636.0
6579.2
6694.0
6465.0
6519.5
6749.3
6500
6750
Molecular Weight
Peak at 23.0 min
53,231
5+
1316.0
6+
1096.0
7+
940.0
948.0
1087.0
1106.0
1281.0
1304.0
1327.0
1338.0
1587.0
6580.0
6408.0
6465.0
6522.0
6635.0
6500
6750
Molecular Weight
1000
1250
1500
m/z

Figure 13


4.4
6.4
4.9
11.5
12.5
13.3
15.7
16.3
100
75
50
25
0
Relative Intensity (%)
0.0
1
5.0
60
10.0
119
15.0
178
20.0
236
25.0
295
30.0
354
Time (min)/Scan

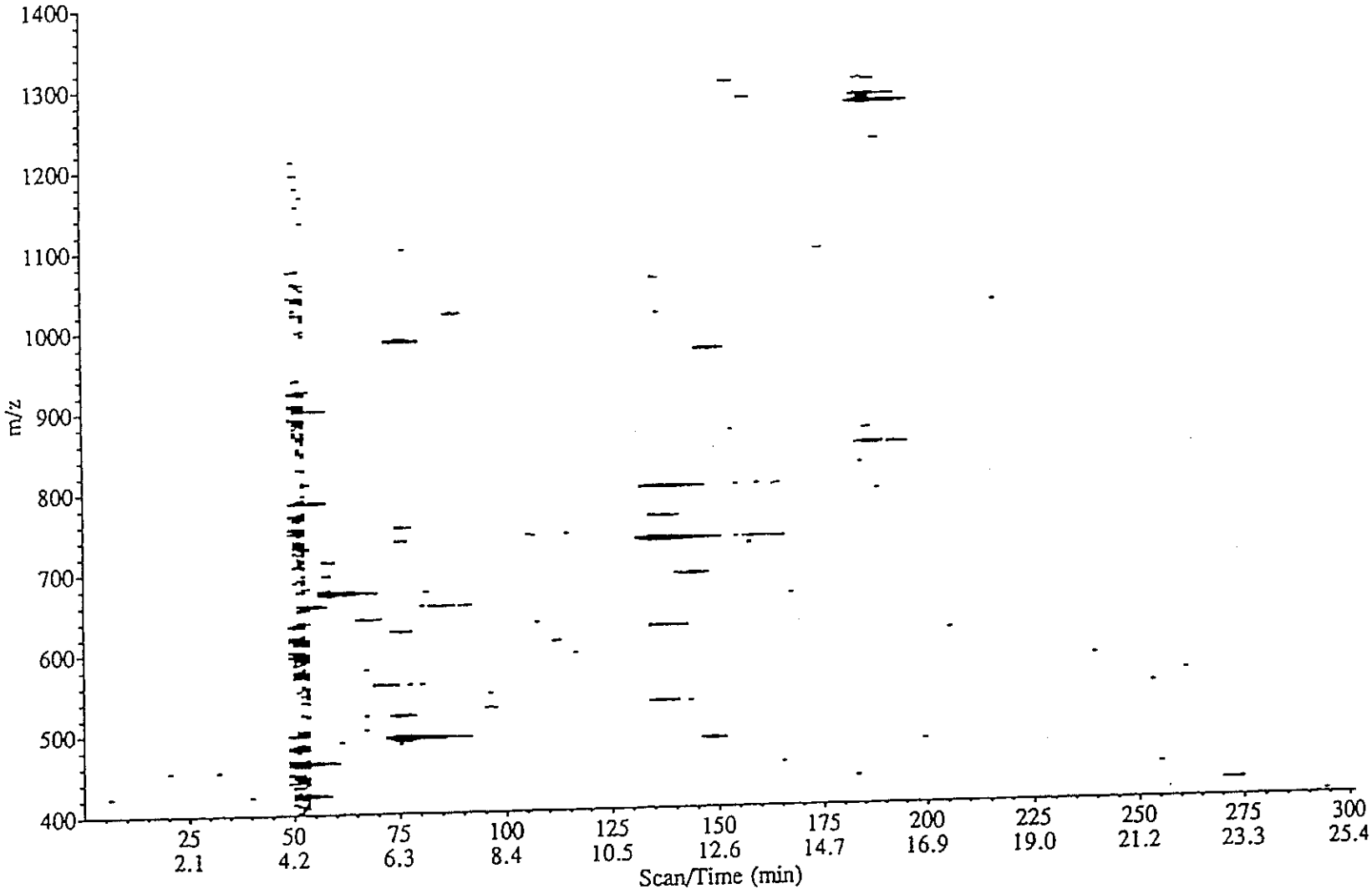

1400
1300
1200
1100
1000
900
800
700
600
500
400
m/z
25
2.1
50
4.2
75
6.3
100
8.4
125
10.5
150
12.6
175
14.7
200
16.9
225
19.0
250
21.2
275
23.3
300
25.4
Scan/Time (min)

Figure 14

Figure 15

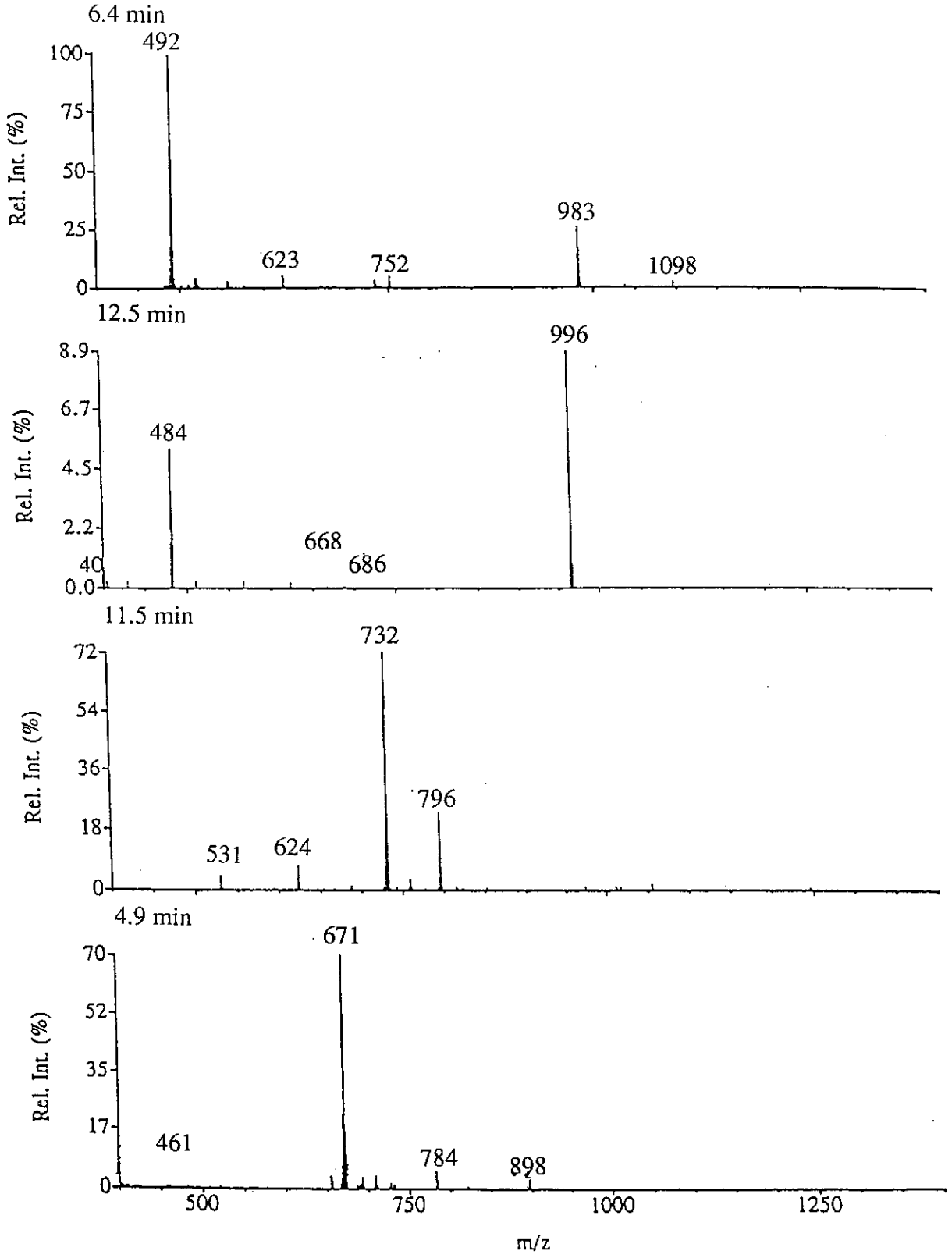

6.4 min
492
100
75
50
25
0
Rel. Int. (%)
623
752
983
1098
12.5 min
996
8.9
6.7
4.5
2.2
0.0
484
668
686
11.5 min
732
72
54
36
18
0
796
531
624
4.9 min
671
70
52
35
17
0
461
784
898
500
750
1000
1250
m/z

Figure 16

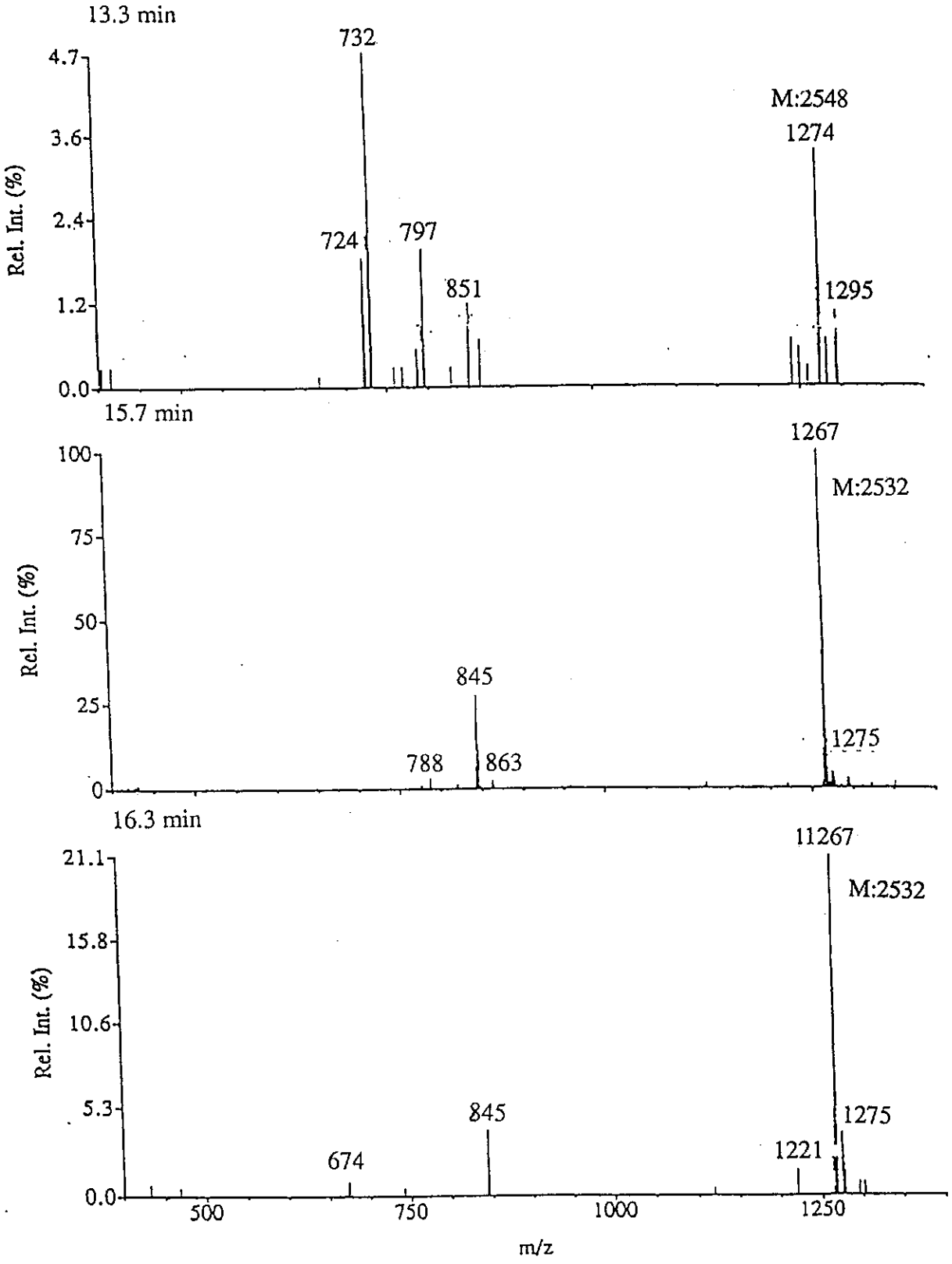

13.3 min
732
M:2548
1274
724
797
851
1295
Rel. Int. (%)
15.7 min
1267
M:2532
845
1275
788
863
16.3 min
11267
M:2532
845
1275
674
1221
m/z

Figure 17

Scylla MT-1

```
PGPCCNDKCVCKEGGCKEGCQCTSCRCSPCEKCSSGCKCANKEECSKTCSKACSCCPT
----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
         10        20        30        40        50        60        70        80
```

N-Terminal Group : Hydrogen

C-Terminal Group : Free Acid

MH+ Monoisotopic Mass = 5997.2577 amu
MH+ Average Mass = 6002.0995 amu
Isoelectric Point (pI) = 7.7

HPLC Index = -236.10
Bull & Breese value = 24410

Elemental composition: C 224 H 379 N 72 O 84 S 18

User-Defined Amino Acid Residues:
None

For Peptide 'Scylla MT-1', Tryptic fragments are:

Num.	Fragment	MH+ (mass)	Sequence
7	39 - 42	435.20	(K) CANK (E)
9	48 - 51	438.20	(K) TCSK (A)
2	9 - 12	452.20	(K) CVCK (E)
3	13 - 17	493.21	(K) EGGCK (E)
6	33 - 38	584.22	(K) CSSGCK (C)
8	43 - 47	595.24	(K) EECSK (T)
5	27 - 32	666.26	(R) CSPCEK (C)
10	52 - 58	684.22	(K) ACSCCPT ( )
1	1 - 8	833.33	( ) PGPCCNDK (C)
2, 3	9 - 17	926.39	(K) CVCKEGGCK (E)
4	18 - 26	986.35	(K) EGCQCTSCR (C)
6, 7	33 - 42	1000.40	(K) CSSGCKCANK (E)
7, 8	39 - 47	1011.42	(K) CANKEECSK (T)
8, 9	43 - 51	1014.42	(K) EECSKTCSK (A)
9, 10	48 - 58	1103.40	(K) TCSKACSCCPT ( )
5, 6	27 - 38	1231.46	(R) CSPCEKCSSGCK (C)
1, 2	1 - 12	1266.51	( ) PGPCCNDKCVCK (E)
3, 4	13 - 26	1460.54	(K) EGGCKEGCQCTSCR (C)
4, 5	18 - 32	1633.59	(K) EGCQCTSCRCSPCEK (C)

Figure 18

Scylla MT-1

```
PGPBBNDKBVBKEGGBKEGBQBTSBRBSPBEKBSSGBKBANKEEBSKTBSKABSBBPT
----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
         10        20        30        40        50        60        70        80
```

N-Terminal Group : Hydrogen

C-Terminal Group : Free Acid

MH+ Monoisotopic Mass = 7023.6428 amu
MH+ Average Mass = 7028.9744 amu
Isoelectric Point (pI) = 10.4

HPLC index = -70.50
Bull & Breese value = 17930

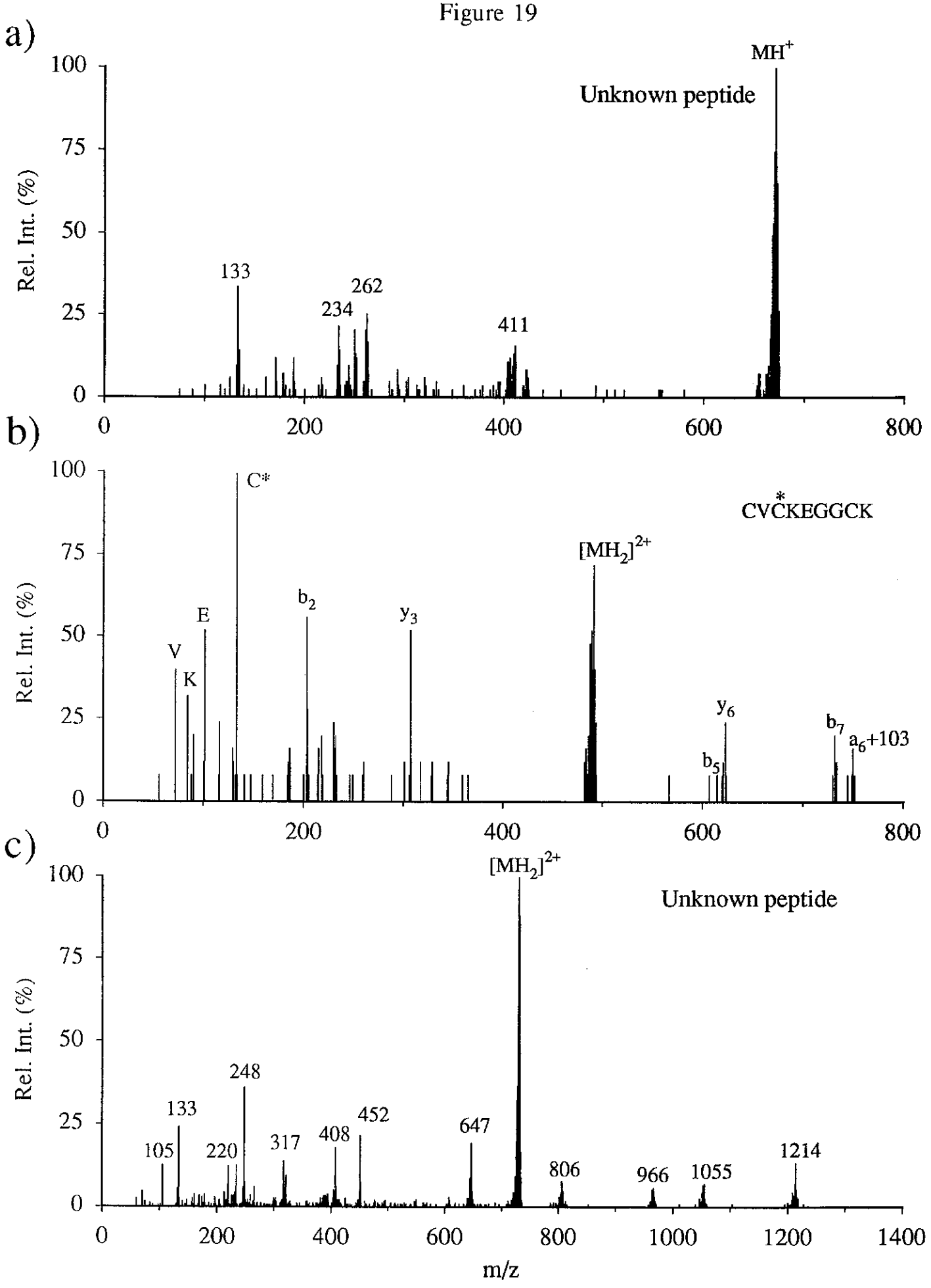
Elemental composition: C 260 H 433 N 90 O 102 S 18

User-Defined Amino Acid Residues:
B - CM-CYS,IODOACETAMIDE

For Peptide 'Scylla MT-1', Tryptic fragments are:

Num.	Fragment	MH+ (mass)	Sequence
7	39 - 42	492.22	(K) BANK (E)
9	48 - 51	495.22	(K) TBSK (A)
3	13 - 17	550.23	(K) EGGBK (E)
2	9 - 12	566.24	(K) BVBK (E)
8	43 - 47	652.26	(K) EEBSK (T)
6	33 - 38	698.26	(K) BSSGBK (B)
5	27 - 32	780.30	(R) BSPBEK (B)
10	52 - 58	855.28	(K) ABSBBPT ( )
1	1 - 8	947.37	( ) PGPBBNDK (B)
2, 3	9 - 17	1097.45	(K) BVBKEGGBK (E)
7, 8	39 - 47	1125.47	(K) BANKEEBSK (T)
8, 9	43 - 51	1128.47	(K) EEBSKTBSK (A)
4	18 - 26	1157.41	(K) EGBQBTSBR (B)
6, 7	33 - 42	1171.47	(K) BSSGBKBANK (E)
9, 10	48 - 58	1331.48	(K) TBSKABSBBPT ( )
5, 6	27 - 38	1459.54	(R) BSPBEKBSSGBK (B)
1, 2	1 - 12	1494.60	( ) PGPBBNDKBVBK (E)
3, 4	13 - 26	1688.62	(K) EGGBKEGBQBTSBR (B)
4, 5	18 - 32	1918.70	(K) EGBQBTSBRBSPBEK (B)

Figure 19


a)
Unknown peptide
MH+
Rel. Int. (%)
133
262
234
411
b)
CVCKEGGCK
C*
[MH2]2+
b2
y3
E
V
K
y6
b7
a6+103
b5
c)
Unknown peptide
[MH2]2+
248
133
105
220
317
408
452
647
806
966
1055
1214
m/z