TANDEM MASS SPECTROMETRIC STUDIES OF PORPHYRINS:
STRUCTURAL AND PHOTOCHEMICAL STUDIES

By

JOHN D. LAYCOCK

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1994
In loving memory of Robert Cole Laycock.
ACKNOWLEDGMENTS

There are many to whom I am greatly indebted for helping me to find the inspiration, gumption, and courage to complete this dissertation. Many are not included here; they know who they are -- I wish them the best.

I want to first thank my late grandfather, Robert C. Laycock, to whom this thesis has been dedicated. He was always interested in understanding what I was learning in my studies of science and chemistry. He gave me much encouragement during my first year of graduate school. I miss our days of sailing his boat out in Pensacola Bay.

I wish to express my deepest gratitude to my parents, Bob and Brigitte, for constantly providing me with loving guidance, wise advice, and helping hands throughout my studies. They have always been there when I needed them; without their help, I doubt I could have completed this dissertation. I wish to thank my sister, Maren, for talking me through my hardest times; in many ways our spiritual growths have coincided. I am sure we will continue to grow together for many years. I thank the rest of my extended family for their gracious encouragement and support.

I want to thank my fellow (former) graduate students Drs. John Pike and Nathan Yates for inspiring me to find the meaning behind putting forth a
consistent work effort. On the other hand, I want to thank John for knowing when it is time to let the beer flow and Nate for knowing when it is time to go surfing. I hope to continue our friendships for many years to come.

I would like to express my gratitude to Dr. Richard A. Yost for his support and professional guidance (and understanding) throughout my graduate studies. I thank Dr. J. Martin E. Quirke for our energetic conversations on porphyrin chemistry during our many "porphyrin weekends"; I admire his boundless enthusiasm for porphyrin research. I thank Dr. Jodie Johnson for training me in the use of the TSQ 45 and for always being willing to help with troubleshooting and other experimental problems.

I would like to thank the Yost Group members (old and new) for their friendship and for their helpful advice with many of the aspects of preparing this dissertation. Finally, I wish to acknowledge the members of my committee, Drs. Willard Harrison, James Winefordner, William Weltner, and Anthony Randazzo.
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS** ........................................ iii

**ABSTRACT** .................................................... viii

**CHAPTERS**

1  **INTRODUCTION** ........................................... 1

   Tandem Mass Spectrometry .................................. 2
   Porphyrins .................................................. 6
      Historical Perspective .................................. 7
      Inorganic Chemistry of Porphyrins .................... 9
      Biological Chemistry of Porphyrins ................... 12
      Geological Chemistry of Porphyrins .................. 15
      Porphyrins of Biomedical Importance ................. 18

   Mass Spectrometry of Porphyrins ....................... 19
      Electron Ionization (EI) ................................ 20
      Chemical Ionization (CI) ............................... 24
      Tandem Mass Spectrometry (MS/MS) .................... 25

   Scope of Dissertation ..................................... 26

2  **CHARACTERIZATION OF CYCLOALKANOPORPHYRIN SKELETAL TYPES USING ELECTRON IONIZATION TANDEM MASS SPECTROMETRY: IMPLICATIONS FOR ANALYSIS OF GEOPORPHYRIN MIXTURES** ........................................ 29

   Introduction ............................................... 29
      Geoporphyrin Skeletal Types .......................... 30
      Analysis of Geoporphyrins ............................. 33

   Experimental .............................................. 35
      Samples ................................................. 35
      Methods ................................................. 35
      Metal Insertion ........................................ 36

   Results and Discussion ................................... 37

   Comparison with the Daughter Ion Spectra of Molecular Ions of Geoporphyrins ......................... 47

   Conclusions ............................................... 55
3 ELECTRON IONIZATION TANDEM MASS SPECTROMETRY OF 5-NO₂ OCTAETHYLPORPHYRIN (OEP): UNUSUAL FRAGMENTATIONS AND THE INFLUENCE OF CHELATED DIVALENT METAL IONS

Introduction ............................................................................. 57
  Spectroscopic Studies of Nitroporphyrins ................................. 57
  Mass Spectrometry of Nitroporphyrins .................................... 59
Experimental ............................................................................ 62
  Compound Preparation ............................................................. 62
    Synthesis of NO₂ porphyrins .................................................. 62
    Synthesis of isotopically labelled porphyrins ......................... 62
    Synthesis of Ni(II) and Co(II) 5-NO₂ OEP ............................. 63
    Synthesis of Zn(II) and Cu(II) 5-NO₂ OEP ............................ 63
    Synthesis of Mg(II) 5-NO₂ OEP ........................................... 65
    Synthesis of Ag(II) 5-NO₂ OEP .......................................... 65
    Synthesis of Pd(II) 5-NO₂ OEP .......................................... 66
Methods .................................................................................. 66
Results and Discussion ............................................................. 67
  Analysis of Free-base 5-NO₂ OEP .......................................... 67
  Analysis of the Divalent Metal Complexes of 5-NO₂ OEP ........... 81
    Spectroscopic studies ......................................................... 81
    Mass Spectrometric studies ............................................... 83
Conclusions ............................................................................. 99

4 TANDEM MASS SPECTROMETRIC STUDIES OF PHOTOSENSITIZING REACTIONS OF PORPHYRINS

Introduction ............................................................................. 102
  Photochemistry ................................................................... 103
  Photosensitizing Reactions .................................................... 103
  Porphyrin Photosensitizers .................................................... 106
  Clinical Photodynamic Therapy (PDT) .................................... 108
  On-line Solution Chemistry .................................................. 109
Thermospray Ionization ............................................................ 112
  Instrumentation ................................................................... 112
  Theory .................................................................................. 115
On-line Flow Cell Development .................................................. 116
  Commercially Available Flow Cells ....................................... 116
  Residence Time Requirements .............................................. 117
# High Pressure Photochemical Reaction Cell

118

## Experimental

120

- On-line Setup
- Off-line Setup
- Compounds
- Methods

## Results and Discussion

129

- Initial TSP/MS of Porphyrins
- On-line
- Off-line
- Conclusions

## 5 CONCLUSIONS AND FUTURE WORK

153

- Cycloalkanoporphyrins
- 5-NO$_2$ OEP and its Divalent Metal Complexes
- Porphyrin Photosensitizer

## REFERENCES

157

## BIOGRAPHICAL SKETCH

168
Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

TANDEM MASS SPECTROMETRIC STUDIES OF PORPHYRINS: STRUCTURAL AND PHOTOCHEMICAL STUDIES

By

John D. Laycock

December 1994

Chairperson: Dr. Richard A. Yost
Major Department: Chemistry

The research reported in this dissertation employs tandem mass spectrometry (MS/MS) to study a variety of topics in porphyrin chemistry. Closely related structures of geoporphyrins containing exocyclic rings are shown to be distinguishable by MS/MS. We show strong evidence that the non-planarity of 5-NO₂ octaethylporphyrin (OEP) and its divalent metal complexes in conjunction with oxygen migration, induces unusual porphyrin fragmentations in the gas phase. Finally, porphyrin photosensitizing reactions are studied by on-line and off-line thermospray ionization (TSP) MS/MS.

The electron ionization tandem mass spectrometric (EIMS/MS) analysis of 22 cycloalkanoporphyrin (CAP) standards is presented. The daughter spectra of molecular ions of 6 different skeletal types show that CAPs bearing different
isocyclic rings can be distinguished by the variation of the relative abundance of three ions ([M-43]⁺, [M-44]⁺, and [M-45]⁺). When the normalized intensities of the three ions are plotted on a ternary diagram, distinct clusters of points appear for each of the skeletal types studied. This approach is applied to daughter spectra of molecular ions from partially separated mixtures of geoporphyrin extracts and the implications for biomarker analysis are discussed.

The EIMS/MS and high resolution EIMS data for free-base 5-NO₂ OEP show that the porphyrin macrocycle molecular ion cleaves a pyrrole unit in the gas phase. This phenomenon is significant since ring scission has previously only been observed for substituted porphyrins by surfaced-induced decomposition during H₂ chemical ionization MS. Daughter spectra for the molecular ions of divalent metal complexes of 5-NO₂ OEP show a metal-dependency in their fragmentations. The most metal-dependent daughter ion, the [M-86]⁺ ion, correlated well with metal- and conformation-sensitive infrared and Raman bands. Oxygen migration and non-planarity are important factors in inducing both of these fragmentations.

Porphyrin photosensitizing reactions of interest to photodynamic therapy for cancer and their characterization by on-line and off-line photochemistry TSP/MS/MS is discussed. The on-line monitoring of the photosensitization of tryptophan by hematoporphyrin is demonstrated. The construction of an on-line photochemistry cell and its evaluation by comparison to off-line photochemistry experiments is discussed.
CHAPTER 1
INTRODUCTION

This dissertation can be divided into two sections (each describing the use of tandem mass spectrometric techniques to study porphyrins). The first section involves structural studies of geoporphyrin standards and synthetic nitroporphyrins. The second involves the structural elucidation of photoproducts of porphyrin photosensitizing reactions. Introductory topics that follow in this chapter are intended to familiarize the reader with background and previous work of relevance to research discussed throughout the dissertation. A brief introduction to tandem mass spectrometric instrumentation and techniques is given. Then a historical perspective of porphyrin research is presented, followed by sections pertaining to the following areas of interest in porphyrin chemistry: inorganic chemistry of porphyrins, biological chemistry of porphyrins, geological chemistry of porphyrins, porphyrins of biomedical significance, and finally, the mass spectrometry of porphyrins.

Background information that is pertinent to each individual chapter will be covered therein. The dissertation will be summarized and put into perspective with previous work at the end of this chapter (Scope of Dissertation).
Tandem Mass Spectrometry

All of the tandem mass spectrometric studies described within this dissertation were performed using a triple quadrupole mass spectrometer (TQMS), (Figure 1-1). The TQMS was first described by Yost and Enke in 1979. The arrangement of the three quadrupoles (Q1, Q2, and Q3) allows for two separate stages of mass analysis (MS/MS) [Yost and Enke, 1979; Yost and Fetterolf, 1983]. The advantages that an additional stage of mass analysis afford are numerous [Johnson and Yost, 1985]. They include: first, a reduction in the chemical noise and thus decreased limits of detection, and second, minimized (or eliminated) need for sample clean-up and/or chromatographic separation.

Figure 1-2a shows a TQMS in the MS operation mode. For single-stage mass separation experiments, either Q1 or Q3 may be scanned. The other two quadrupoles are operated in the radio frequency (RF) only mode, allowing essentially all masses to pass. The two MS/MS scanning modes that were utilized in this dissertation research were the daughter scan and the parent scan.

In the daughter scan experiment (Figure 1-2b), one ion of a particular mass-to-charge ratio (m/z) of interest (the parent ion) is allowed to pass through the first quadrupole (Q1). This ion is then passed through the second quadrupole with a selected collision energy. Within Q2 (RF only), the selected ion undergoes collisions with an inert target gas (typically 1-3 mtorr of nitrogen or argon). The collision imparts internal energy to the analyte ion; if the energy is sufficient,
Figure 1.1: The triple quadrupole mass spectrometer
Figure 1-2: Scan modes of the triple quadrupole mass spectrometer: a) Q1 MS; b) daughter scan; c) parent scan; d) selected reaction monitoring SRM scan.
fragmentation of the ion, or collisionally induced dissociation (CID), occurs. The resultant fragment ions (daughter ions) pass into the third quadrupole Q3 and they can be mass analyzed. The spectrum thus obtained by scanning Q3 is called a daughter spectrum. The daughter spectrum is particularly useful in the structural elucidation of analyte ions.

In the parent scan (Figure 1-2c), a range of parent ions are allowed to pass through Q1. These ions are then passed through the collision cell (Q2), where CID occurs. The Q3 is set such that only one daughter ion of a particular m/z is passed. Any ions which fragment to give a daughter ion of this m/z will be detected and are referred to as parent ions. The parent spectrum is useful in identifying ions which fragment to form a common substructure. The parent scan is also useful in determining mechanisms of the reaction pathways of a given parent ion.

The various MS/MS scanning modes allow for the development of unique analytical schemes. For instance, selected reactions can be monitored (SRM) by passing only one parent ion through Q1 and passing only selected daughter ions through Q3 (Figure 1-2d). The chemical noise is significantly reduced in the SRM mode and trace compounds can be detected in complex matrices [Johnson and Yost, 1985; Johnson et al., 1986].

The SRM scheme may be also used to study ion/molecule reactions. By replacing the inert collision gas of Q2 with a reactive gas, Q2 may be employed as a reaction chamber. The scanning mode is essentially the same as the SRM
scan described above except that Q3 is scanned to allow reaction products to pass. For instance, Freeman et al. [1990b] developed a method for reacting model nucleophile or DNA base ions with electrophilic gases; Q3 was scanned to detect adduct ions. The utilization of this scheme in order to mimic DNA/carcinogen reactions in solution has been shown [Freeman et al., 1990b; Annachino, 1993].

**Porphyrons**

Porphyrons are cyclic tetapyrroles that are of great importance to all living cells; they are the active structural moiety of hemoglobin, myoglobin, all the cytochromes, and the chlorophylls (dihydroporphyrins). The structure of porphine, the simplest porphyrin, is shown in Figure 1-3. Four pyrrole units are linked by four one-carbon bridges; the entire structure has a planar geometry due to the conjugation of the sp² carbons. The bridge carbons are termed the meso positions. The pyrrolic carbons adjacent the nitrogen are α-positions and the peripheral carbons are β-positions. The IUPAC numbering system is also shown in Figure 1-3. The porphyrin may be substituted by a variety of functional groups at either the meso or β positions.
Figure 1-3: Porphine, the simplest porphyrin.
In 1844, Verdeil chemically altered chlorophyll into a red pigment and suggested that a structural relationship exists between chlorophyll and heme [Dolphin, 1978]. In 1915, Willstatter received the Nobel Prize for his purification and structural determination of plant pigments, especially that of the chlorophylls [James, 1993]. About a decade later, Hans Fischer, widely regarded as the "father" of modern porphyrin chemistry, was awarded a Nobel Prize for his synthesis of hemin from dipyrrromethanes. Fischer later reviewed the early work of porphyrin structural and synthetic studies in his three volume work Die Chemie des Pyrrols [Fischer and Orth 1937a,b; Fischer and Stern 1940]. In 1934 Alfred Treibs, a colleague of Fischer's, identified vanadyl deoxophyloerythroetioporphyrin (DPEP) as a major metalloporphyrin in petroleums and Shale. He proposed ancient chlorophylls as the source of DPEP; out of this original theory, the field of molecular organic geochemistry developed (see Figure 1-6 and the section on the geochemistry of porphyrins).

In 1962 Kendrew and Perutz shared the Nobel Prize in chemistry for their structural studies on the oxygen-binding proteins hemoglobin and myoglobin by X-ray crystallography (Watson and Crick) shared the Nobel Prize in medicine during the same year for their similar structural elucidation of DNA) [James, 1993]. Robert Woodward, considered by many the foremost organic chemist of the
twentieth century, was awarded the 1965 Nobel Prize for his total synthesis of the
tetrapyrrole vitamin B$_{12}$ [James, 1993].

In 1964, Falk published his book *Porphyrrins and Metalloporphyrins* which
discussed and reviewed some of the physico-chemical aspects of porphyrins.
However, most of the porphyrin research until the mid 1970's dealt with structural,
synthetic and biosynthetic topics. More recent research focuses on the function
and mode of action of porphyrins in the various important biological functions.
Dolphin edited a seven-volume series covering a wide range of topics in the
chemistry and biochemistry of porphyrins in 1978.

Research interest in porphyrins has increased dramatically in recent years
as scientists are discovering a diverse array of reactions dependent on porphyrins
for their function. Much of the present porphyrin research is aimed at
understanding the fundamental processes of photosynthesis and catalysis.
Deisenhofer, Huber and Michel won a Nobel Prize in 1988 for their work probing
the sequence of events occurring in porphyrin chromophores utilizing magnetic
resonance techniques, laser spectroscopies, and X-ray crystallography
[Deisenhofer, 1989]. With *in vivo* functionality and electronic properties of
porphyrins better understood, *in vitro* biomimetic photocatalysts may be
eventually developed [Fajer, 1991]. Porphyrins are also of current interest in the
area of molecular electronics; Crossley and Burn [1991] have recently synthesized
a "wire" consisting of a linear array of porphyrin macrocycles.
Inorganic Chemistry of Porphyrins

Porphyrins chelate with virtually all metallic elements. The coordination of a metal significantly alters the physical, chemical and spectroscopic properties of the porphyrin. The free-base porphyrin $H_2(P)$ chelates with the metal (M) to form the metalloporphyrin $M(P)$ in what is termed the metallation reaction (Figure 1-4). The simplest case, and the only one that will be dealt with here, is the formation of a square-planar complex from a dipositive metal ion ($M^{2+}$) and a dinegative porphyrin anion ($P^2$). The reverse of this reaction is demetallation, and occurs in the presence of acids (Figure 1-4).

Five different stability classes of metalloporphyrins have been distinguished based on their ease of demetallation in the presence of various protic reagents [Falk 1964; Buchler, 1975]. Buchler used a stability index to relate inherent properties of the metal to the stability of the metalloporphyrin complex [Buchler, 1975; Buchler, 1978]. The Buchler stability index ($S_I$) is easily calculated from fundamental properties of the metal:

$$S_I = 100(PE)(V)/R$$

Where PE is the Pauling electronegativity of the metal, V is the valence of the metal, and R is the ionic radius of the metal. The lower the stability index, the greater the ease of demetallation of the porphyrin in solution.
Figure 1-4: Metallation reaction of porphyrin with divalent metal.

\[ \text{M}^{2+} \xrightleftharpoons{2\text{H}^+} \text{H}_2\text{(P)} \]
Porphyrrins can have a significant loss of planarity with the coordination of certain metals. This is of interest because many biologically active porphyrins are non-planar.

**Biological Chemistry of Porphyrins**

Cyclic tetrapyrroles are involved in many vital biological processes. Electron and oxygen transport processes, photosynthetic reactions, oxygen insertion, and CO$_2$ removal from tissues are all dependent on cyclic tetrapyrroles [Addison et al., 1977]. Four important types of cyclic tetrapyrroles are porphyrins, chlorins (dihydroporphyrins), bacteriochlorins (tetrahydroporphyrins), and corrins.

In Figure 1-5, the structure of a metalloporphyrin and a metallochlorin are shown. In the chlorin, a peripheral double bond has been removed through reduction; thus, two additional hydrogens are present at the $\beta$-carbons of the pyrrole ring. The mere addition of these two hydrogens is responsible for a substantial shift in the UV-Vis absorption peaks. Thus, while most porphyrins are red in color, the chlorins are green.

Also shown in Figure 1-5 is the structure of protoheme (heme). Heme is the iron (II) complex of protoporphyrin IX, and is the prosthetic group in hemoglobin, myoglobin, the cytochromes, and many of the peroxidases and catalases. The iron (II) ion is usually octahedrally coordinated, meaning that two additional ligation sites remain available in the heme-porphyrin complex.
Figure 1-5: Structures of porphyrins and chlorins.
Hemoglobin contains four protein-bound heme moieties, and is vital in the transportation of oxygen in the bloodstream. The oxygen is carried by axial (perpendicular to the ring) ligation to the iron. Once one molecule of oxygen is bound, the other coordination sites of the hemoglobin have an increased affinity for binding oxygen [Marks, 1969].

In deoxymyoglobin, one of the axial sites of the protein-bound heme complex is taken up by a nitrogen of histidine. In myoglobin, an oxygen is reversibly bound at the other axial position. This oxygen is only relinquished when muscle oxygen demand is greater than the oxygen supply from the blood (during strenuous exercise). In this way, myoglobin acts as an emergency store of oxygen [Marks, 1969].

The cytochromes are responsible for the respiratory electron transport processes occurring in the mitochondria and the chloroplasts. Cytochromes are classified as a, b, or c depending on their characteristic absorption spectrum, with cytochrome b, c, and c, having heme as its prosthetic group. The cytochromes transfer electrons through reversible changes in the iron atom between the reduced iron (II) and oxidized iron (III) states [Lemberg and Barrett, 1973; Ortiz de Montellano, 1986].

The structure of chlorophyll a is shown in Figure 1-5. The structure is a chlorin, as one of the peripheral double bonds is absent (thus the green color). The coordinated metal is magnesium. The chlorophylls are responsible for the photosynthetic reactions that harvest the sun's energy and store it as chemical
potential energy. Bacteriochlorophylls constitute the light-harvesting antenules of photosynthetic bacteria [Deisenhofer and Michel, 1989; Tronrud et al., 1986]. The bacteria funnel photons into a complex called the reaction center; a separation of oppositely charged ions across a membrane creates a potential gradient. The stored electrical energy is the driving force of the biochemistry of the organism [Norris and Schiffer, 1990; Feher et al., 1989].

Vitamin B$_{12}$, a corrin, is a cyclic tetrapyrrole structurally related to heme, but with a cobalt (II) ion as the coordinated metal. Vitamin B$_{12}$ is an essential nutrient and is present in coenzymes responsible for a number of reactions, including the oxidation of fatty acids and the synthesis of DNA [Addison et al., 1977].

In many enzymes a metalloporphyrin is the biochemically active site. Porphyrins are of great interest in the development of new catalysts; much research has been devoted to the study of porphyrin-initiated reactions. Some porphyrins are known to be strong photosensitizing agents. Because of these photosensitizing properties, porphyrins have found utility as photodynamic therapy (PDT) agents in the treatment of cancer [Doiron and Gomer, 1983]. Details on PDT are covered in detail in Chapter 4 of this dissertation.

Geological Chemistry of Porphyrins

The inherent stability of the porphyrin macrocycle coupled with the abundance of porphyrinic structures in living systems makes them important to
geochemistry. Treibs [1934; 1936] hypothesized that geoporphyrins derived from the chlorophyll in ancient plants undergo decarboxylation, reduction, dealkylation, and oxidation. Figure 1-6 shows the series of reactions occurring in the diagenesis scheme proposed by Treibs [1934; 1936]. The color of each compound and its nominal molecular weight is shown as these compounds have been classified by UV-Vis spectroscopy [Stern and Wenderlein, 1936] and by mass spectrometry [Baker et al., 1968; Baker and Smith, 1973].

The first reactions involve the demetallation, saponification, and decarbomethoxylation of chlorophyll a; Treibs proposed that these reactions occur at the water/sediment interface with little external stress. The subsequent reactions (reduction, aromatization, decarboxylation, and chelation) are said to occur as the sediments are compacted and the thermal stress increases. Treibs explained that the order of these reactions are interchangeable and are dependent on the depositional environment.

Porphyrins resulting from these reactions are ubiquitous in petroleum deposits, yet comprise only a small portion of the total carbon. There are a variety of peripheral substituents present (meso substitutions generally do not occur except for isocyclic rings) making the geoporphyrins important as biomarkers [Eglinton and Calvin, 1967].

Biomarkers are defined as compounds which retain a basic carbon skeleton during and after deposition in the earth's sediments. The biomarkers can be related to a specific source of organic materials and for this reason are
Chlorophyll $a$
- bright green
- Mol. wt. 892

vanadyl DPEP
- orange pink
- Mol. wt. 541

demetallation
- saponification
- decarbmethoxylation

yellow green
- Mol. wt. 534

reduction

olive green
- Mol. wt. 522

aromatization,
- chlorin-porphyrin
- transition

red brown
- Mol. wt. 476

decarboxylation

red brown
- Mol. wt. 520

Figure 1-6: The Treibs scheme.
sometimes referred to as chemical fossils [Bonnett et al., 1991]. Geoporphyrins have proven useful as biomarkers because their structural complexity allows for fingerprinting of samples for oil-oil or oil-source rock correlations [Corwin, 1960; Callot et al., 1987; Chicarelli et al., 1987; Filby and Van Berkel, 1987]. Current research aims at clarifying the evolution of biological precursors into geoporphyrin products. That extended isocyclic rings (up to seven carbons) and benzoporphyrins are present in geoporphyrin extracts suggests that some processes are unaccounted for under Treibs' digenesis schemes [Bonnett et al., 1991]. The Treibs scheme also does not account for the occurrence of high carbon-number porphyrins observed by Johnson et al. [1985] by tandem mass spectrometry. Some of the previous porphyrin work in our research group has been aimed at obtaining structural information on geoporphyrins using tandem mass spectrometry [Quirke et al., 1988; Beato et al., 1989a,b; Quirke et al., 1989].

Porphyrrins of Biomedical Importance

Porphyrrins are of great importance to many areas of the biomedical and pharmaceutical sciences [Doiron and Gomer, 1984; Bonnet et al., 1989; Gomer, 1990; Morris, 1991; Kessel et al., 1991], especially those involving photoirradiation. It has long been known that certain porphyrrins cause photosensitivity in man. The photosensitivity associated with malfunctions of
porphyrin metabolism, such as in the porphirias, is attributed to the presence of photoactive porphyrins under the skin [Blum, 1941].

The use of photosensitizing porphyrins in the treatment of malignant tumors is becoming increasingly important. Thousands of cancer patients have been treated with Photofrin I, a commercial product that is employed for photodynamic therapy (PDT) [Gomer, 1990]. Photofrin consists of a complex mixture of porphyrin polymers derived from hematoporphyrin. One of the major current research efforts centers on characterizing the Photofrin mixture in order to determine which of the its components are the active PDT agents [Bonnett et al., 1981; Berenbaum et al., 1982; Dougherty, 1984]. Another well-researched area involves the development of a new generation of synthetic cationic and anionic water-soluble porphyrin PDT agents [Le Nouen et al., 1989; Pasternack et al., 1985]. Despite considerable research efforts in these areas, the modes of action of the porphyrins are not well known. It is uncertain how cell components (and which of them) sustain enough damage to cause cell death.

Mass Spectrometry of Porphyrins

Mass spectrometry has been shown to be a powerful tool for the structure elucidation and identification of porphyrins. Past research in our group has been involved with the detection and structure elucidation of high-carbon number geoporphyrins [Johnson et al., 1986; Cuesta et al., 1989] and with the
investigation of porphyrin related phenomena. Tandem mass spectrometry of doubly charged ions and the surface-induced decomposition of porphyrins in chemical ionization are two areas that have been well-researched [Beato et al., 1989a; Beato et al., 1991].

Electron Ionization (EI)

A compound that serves as a model for the fragmentation of biological and geochemical porphyrins is octaethylporphyrin (OEP). The full-scan EI mass spectrum of OEP is shown in Figure 1-7a. The ring structure remains intact as evidenced by the absence of mono-, di-, or tripolylic ions in the EI mass spectrum. The conjugation and high stability of the ring prevent the loss of pyrollic units. In Figure 1-7b, the OEP mass spectrum has been blown up by a factor of five and expanded to emphasize the singly charged region of the mass spectrum. The spectrum shows that eight successive losses of 15 mass units occur from the ions. Seibl first reported in 1968 that the fragmentation of the peripheral ethyl substituents of OEP occurs via a benzylic type β-cleavage in which successive losses of methyl radicals from the molecular ion are observed (Figure 1-8). Regardless of the length of the alkyl chain, the fragmentation of alkyl-substituted porphyrins is of the β type [Britton, 1985; Johnson et al., 1986]. The bond between the first and second carbons is homolytically cleaved to form an alkyl
Figure 1-7: Electron ionization mass spectrum of OEP: a) Full-scan spectrum showing the singly and doubly charged region of the mass spectrum; b) Expanded and blown up view of the spectrum showing the singly charged region of the mass spectrum.
Figure 1-8: Fragmentation of OEP with the characteristic loss of methyl radicals.
radical. This phenomenon has proven useful for the determination of the peripheral substituents of porphyrins.

There is a high abundance of doubly charged ions in the mass spectra of porphyrins (Figure 1-7a); this can be attributed to the large number of $\pi$ electrons that are readily available [Jackson et al., 1965]. Nuclear magnetic resonance (NMR) data show that the charges are separated by localization at two of the central nitrogens [Chakraborty et al., 1982]. The $\beta$-cleavage fragmentation occurring in the doubly charged region of the mass spectrum is significantly greater than is observed in the singly charged region. Additionally, some $\alpha$-cleavage occurs in the doubly charged region [Beato et al., 1989b].

There is a marked influence on the fragmentation that is dependent on the chelated metal ion. The relative abundance of the ions in the doubly charged region of the mass spectrum has been shown to correlate with the modified Buchler stability index ($S_\nu$) for the metal, while the singly charged region remains relatively unaffected by the nature of the metal [Beato et al., 1989b].

**Chemical Ionization (CI)**

Chemical Ionization with methane as the reagent gas has proven useful for the determination of porphyrin molecular weights [Eglinton et al., 1979]; however, the most interesting phenomenon occurring in chemical ionization of porphyrins is surface-induced decomposition in the presence of $\text{H}_2$ or $\text{NH}_3$ [Beato et al., 1991]. Porphyrin molecules that have deposited on the inside of the ion volume...
are reduced and then subsequently revaporized as a result of radiative heating by the direct exposure probe. As a result of the reduction to porphyrinogen (a hexahydroporphyrin), the bridge carbons are now sp$^3$ hybridized; thus they are no longer conjugated. This acid-labile macrocycle fragments to give mono, di, and tripyrrole fragments which yield structural information with regards to pyrrole sequence and individual pyrrole composition [Beato et al., 1991; Van Berkel et al. 1989a,b; Shaw et al., 1981].

**Tandem Mass Spectrometry (MS/MS)**

Tandem mass spectrometry has been utilized to gain structural information from the molecular ions of various porphyrins and especially those of geoporphyrins [Johnson et al., 1986]. Partially separated mixtures of geoporphyrins can be analyzed by taking daughter spectra of the molecular ions of interest. Structural information on these porphyrins is thus obtained. By performing MS/MS, the nature of the peripheral substituents may be clarified.

Tandem mass spectrometry of doubly charged porphyrin ions is of interest since porphyrins are one of the few classes of compounds (small peptides and polycyclic aromatic hydrocarbons are others) whose ions retain a double charge when fragmented via CID [Appling et al., 1983; Hanner et al., 1982]. Most compounds with doubly charged ions would retain only a single charge after CID owing to the charge exchange reactions that occur in the collision cell (Q2).
Doubly charged porphyrin ions are different in this respect because they are resonance-stabilized and maintain their double charge during CID [Beato et al., 1989b].

Scope of Dissertation

The dissertation is divided into two sections based on the methods employed and the analytical applicability of the work presented. Chapters 2 and 3 concentrate on the study of ion fragmentation pathways of series of porphyrin standards in the gas phase. Chapter 4 is a study of the aqueous-phase photosensitizing reactions of porphyrins with a biological substrate. Chapter 5 summarizes the dissertation and outlines areas of future work.

In Chapter 2, the electron ionization tandem mass spectrometric (El/MS/MS) analysis of cycloalkanoporphyrin (CAP) standards is presented. First, a background of biomarker analysis and a survey of current research in this area is discussed. The structure of the skeletal types studied and their possible relationship to naturally occurring geoporphyrins is shown.

The daughter spectra of the molecular ions of 6 different skeletal types are presented. The data show that the CAPs bearing different isocyclic rings can be distinguished by the variation of the relative abundances of three ions ([M-43]+, [M-44]+, and [M-45]+) within each compound's spectrum. When the three ion's normalized intensities are plotted on a ternary diagram, distinct clusters of points
appear for each of the skeletal types studied. For comparison, the technique is applied to daughter spectra of molecular ions from partially separated mixtures of geoporphyrin extracts. The implications of this work towards porphyrin biomarker analysis is discussed.

In chapter 3, meso-nitro octaethylporphyrin (5-NO₂ OEP) and its divalent metal complexes are studied by electron ionization tandem mass spectrometry (EI/MS/MS) and by high resolution EI mass spectrometry. The reason 5-NO₂ OEP was chosen is that has been studied extensively by various optical spectroscopic techniques. The structure is known to be non-planar; indeed, 5-NO₂ OEP has been shown to be useful as a model compound for cytochrome-c [Shelnutt et al., 1991a,b].

The data for the free base 5-NO₂ OEP show that the porphyrin macrocycle molecular ion cleaves a pyrrole unit in the gas phase. This phenomenon is significant since ring scission has only previously occurred for substituted porphyrins by surfaced-induced decomposition [Beato et al., 1991]. The results of isotopic labelling with ¹³C and ¹⁵N and high resolution mass spectrometry confirm ion compositions deduced from MS/MS spectra of unlabelled 5-NO₂ OEP. The migration of oxygen occurs in an unusual multi-step fragmentation pathway in which several bonds are broken.

Divalent metal complexes of 5-NO₂ OEP are also analyzed by MS/MS. A daughter spectrum is obtained for the molecular ions of each of the various metalloporphyrins. The fragmentations are discussed and compared to those
observed for the free-base 5-NO₂ OEP. Trends within the series of metals studied are discussed and are related to other metal-dependent spectroscopic properties.

In Chapter 4, photosensitizing reactions of porphyrins and their characterization by off-line and on-line photochemistry/thermospray ionization tandem mass spectrometry is discussed. The intended application of this work is to the study of the reactions of photodynamic therapy cancer treatment agents.

The chapter begins with an introduction to photosensitizing reactions and goes on to discuss the utilization of these photosensitizers as therapeutic cancer treatment agents. Some of the techniques for studying on-line solution phase reactions as well as photochemical reactions will be described.

Finally, the development of our instrumentation and methodology is discussed. Comparisons are drawn between our system and commercially available components and between the methods of on-line and off-line experiments. The results of the photolysis experiments are discussed and put into perspective with known biological redox reaction pathways.

Chapter 5 summarizes the work presented in the dissertation. Areas of possible future work are discussed for each of the research topics discussed in the chapters.
CHAPTER 2
CHARACTERIZATION OF CYCLOALKANOPORPHYRIN SKELETAL TYPES USING ELECTRON IONIZATION TANDEM MASS SPECTROMETRY: IMPLICATIONS FOR ANALYSIS OF GEOPORPHYRIN MIXTURES

Introduction

The electron ionization tandem mass spectrometric analyses of six skeletal types of cycloalkanoporphyrins (CAP’s) together with their Ni(II) and VO(II) complexes are presented. Analysis of the daughter ion spectra of the molecular ions \((M^+)\) indicates that it is possible to distinguish isomers with different sizes of isocyclic rings. For instance, deoxophylloerythroetio, DPEP, porphyrins with a 5-membered isocyclic ring show a characteristic, intense \([M-44]^+\) daughter ion, whereas CAP-6 porphyrins (CAPs bearing a 6-membered isocyclic ring) display an intense \([M-45]^+\) daughter ion. As a result of these findings, it is now possible to identify the skeletal type of CAP porphyrins in intact or partially separated geoporphyrin mixtures. Data from previous studies on the porphyrins of New Albany Shale (Mississippian-Devonian, Indiana USA) are used to demonstrate the potential applications of this approach.
Geoporphyrin Skeletal Types

Geoporphyrins (geologically-occurring porphyrins) occur mainly as complicated mixtures of Ni(II) or VO(II) complexes in sediments, oil Shales, bitumens, coals, phosphorites and crude oils [Treibs, 1934]. At least ten different skeletal types of geoporphyrins are known [Chicarelli et al., 1987; Filby and Van Berkel, 1987; Ocampo et al., 1987; Callot et al., 1990; Keely et al., 1990]. In Figure 2-1 the structural types studied here are shown; for the remainder of this chapter the reader will be referred to specific structures within Figure 2-1 by an underlined numeral. Two skeletal types are believed to occur most commonly—the ETIO (1) and the deoxophyloerythroetio (DPEP) (2). The term, CAP-n porphyrin, will be used to describe cycloalkanoporphyrrins bearing a single isocyclic ring with n-carbons (i.e., the DPEP is a CAP-5 porphyrin). Similarly, CAP-N-Me indicates that there is a methyl substituent on the isocyclic ring. Four different skeletal types of CAP-porphyrin have been isolated from geological samples (2,3,4,6; R\(^9\) = CH\(_3\)). The tetrahydrobenzoporphyrrins (THB, 7) themselves have not been reported; however, the di-CAP analogues (8a,b) and benzoporphyrrins have been identified [Kaur et al., 1986; Verne-Mismer et al., 1987]. Thus, the THB skeleton may also occur (as might the des-methyl CAP-6, 5).

The geoporphyrins are a potentially valuable compound class for the fingerprinting of oils in oil-oil and oil-source rock correlation studies in oil
Figure 2-1: Molecular structures of cycloalkanoporphyrin (CAP) standards.
(a) $R_1, R_2, R_3, R_5, R_8 = CH_3$
R_4, R_7 = C_2H_5
(b) $R_1, R_4, R_5, R_8 = CH_3$
R_2, R_3 = C_2H_5; R_7 = H
(c) $R_1, R_4, R_5, R_8 = CH_3$
R_2, R_3, R_7 = C_2H_5

Figure 2-1 -- continued.
exploration [Beato et al., 1989a; Concha et al., 1991]. Similarly, they may prove useful for the identification of oils and tar balls from oil spills [Johnson et al., 1986]. In order to fully realize their potential for such applications, the components of the mixtures must be separated, characterized and quantitated precisely and efficiently.

**Analysis of Geoporphyrins**

Recently, significant advances have been made in both gas chromatographic and high performance liquid chromatographic separations of the metal-free, nickel and vanadyl geoporphyrin mixtures [Sundararaman, 1985; Barwise et al., 1986; Chicarelli et al., 1986; Callot et al., 1990; Peng et al., 1992]. The Bristol [Chicarelli et al., 1987] and Strasbourg [Ocampo et al., 1987; Callot et al., 1987] groups have provided a valuable database of geoporphyrin structures as a result of their elegant structure elucidation studies. The work of the Freeman group on the identification of geoporphyrins using visible spectrophotometry provides intriguing possibilities for rapid characterization [Freeman and Haver, 1990; Freeman et al., 1990a]. High performance liquid chromatography/mass spectrometry (LC/MS) also remains a technique of considerable potential [Bonnet et al., 1991]. In particular, the development of LC/MS using electrospray ionization provides an excellent method for the quantitation of geoporphyrin mixtures [Van Berkel et al., 1993]. Nevertheless, the presence of a range of
isomeric CAP porphyrins in geoporphyrin mixtures provides a formidable barrier to complete analysis of such samples.

The CAP porphyrins may be distinguished by \(^1\)H NMR [Callot et al., 1987; Ocampo et al., 1987]; however, this requires the isolation of individual compounds in substantial quantities (typically >100 \(\mu\)g). This is entirely unsuitable for routine mixture analysis. Clearly, the best hope for success would seem to lie in either GC/MS or LC/MS analyses of the mixtures, but there are no previous reports that it is possible to distinguish and assign isomeric CAP porphyrins by mass spectrometry.

Previous research in our group has demonstrated that electron ionization tandem mass spectrometry (EIMS/MS) is a powerful analytical tool for characterization of geoporphyrin mixtures [Johnson et al., 1986; Quirke et al., 1989; Beato et al., 1991; Concha et al., 1991]. In addition, we reported that many of the CAP geoporphyrins isolated displayed an abundant, unexpected \([M-44]^+\) ion in the daughter ion spectra of their molecular ions, \(M^+\) [Quirke et al., 1989; Beato et al., 1991]. At that time, however, it was impossible to determine whether the isomeric CAP porphyrins could be distinguished by such spectra because pure standards were unavailable.

In this chapter, the EIMS/MS daughter ion spectra of the molecular ions of each skeletal type of CAP geoporphyrin are presented. Then the implications of the data for analysis of geoporphyrin mixtures are discussed.
Experimental

Samples

The Ni(II) and VO(II) complexes of the CAP-5 porphyrins 2a, 3 were gifts from Dr. R. Ocampo (Université Louis Pasteur, Strasbourg). The CAP-6 porphyrins 4a,b, 5a and the CAP-7 porphyrin 6 were gifts from Dr. P. S. Clezy (University of New South Wales, Sidney). The CAP-5 porphyrin 2b, the CAP-6 porphyrins 5b,c and the THB porphyrin 7 were gifts from Dr. T. D. Lash (Illinois State University, Normal). The C3, CAP-5 porphyrin 2c was isolated from the bitumen Gilsonite (Eocene, Uinta Basin, Utah, USA) using the method of Quirke and Maxwell [1980]. The procedures for the isolation and initial characterization of the porphyrins from the New Albany Shale (Mississippian-Devonian, Indiana USA) have been described previously [Beato et al., 1991].

Methods

All mass spectra were obtained using a Finnigan TSQ 45 triple quadrupole tandem mass spectrometer equipped with an INCOS data system. For EIMS, the mass spectrometer was tuned using perfluorotributylamine (FC43). Tuning was not deemed to be complete until there was complete resolution of the ions m/z 502 and 503. EIMS spectra were obtained using a direct exposure probe (DEP).
The porphyrin in dichloromethane (DCM) solution was placed on the DEP filament and allowed to evaporate. The DEP was then heated from ambient to 600°C at 600°C/min. The El emission filament current was 0.3 mA with an electron energy of 70 eV. The spectra were obtained by scanning the first quadrupole, Q1, from m/z 150 - m/z 800 at a rate of 0.8 s. In EIMS/MS mode, the instrument was tuned using Ag(II) 5-nitro-2,3,7,8,12,13,17,18-octaethylporphyrin. We have discovered this porphyrin to be particularly valuable for tuning because the [M-75]+ ion is by far the most abundant daughter ion of the M+ ion; the typical fragmentation pathways by β-cleavage of the ethyl substituents are not observed (see Chapter 3 of this dissertation). The sample was volatilized using the DEP as described previously. The porphyrins were ionized with an electron energy of 70 eV and an emission current of 0.3 mA. The daughter ion spectra were obtained using argon (1.6 mtorr) for collisionally activated dissociation in the second quadrupole with collision energies ranging from 10 eV to 28.5 eV.

Metal Insertion

Nickel porphyrins were prepared by treatment of the metal-free porphyrin with excess Ni(II) acetate in refluxing acetic acid as described by Buchler [1975]. Vanadyl porphyrins were prepared by refluxing the porphyrin with vanadyl sulfate in either acetic acid/pyridine (2:1 v:v) as described by Stanley et al. [1991] or in N,N-dimethylformamide as described by Adler et al. [1970].
Results and Discussion

Where possible the metal-free, Ni(II) and VO(II) complexes of each CAP porphyrin were studied by EIMS/MS; however, sometimes there was insufficient sample to prepare the three compounds. The daughter ion spectra of the molecular ions, M⁺⁺, of all the metal-free, Ni(II) and VO(II) complexes of the CAP porphyrins are summarized in Table 2-1. The daughter ion spectral data for selected porphyrins, CAP-5 [Ni(II) 2a], CAP-5-Me [VO(II) 3] CAP-6 (metal-free 5a) CAP-6-Me (metal-free 4b) CAP-7 [Ni(II) 6] and THB [VO(II) 7] are shown in Figure 2-2.

Previous EIMS/MS analyses of Ni(II) and VO(II) complexes and the corresponding demetallated geoporphyrins from the New Albany Shale (Mississippian-Devonian, New Albany, Indiana, USA), and demetallated VO(II) porphyrins from Boscan oil (Cretaceous, W. Venezuela) revealed that the daughter ion spectra of the molecular ions were quite reproducible over long time periods (ten years) and with different operators of the instrument [Johnson et al., 1986; Quirke et al., 1989; Beato et al., 1991; Concha et al., 1991]. Nevertheless, it was essential to confirm the reproducibility of the method using CAP-standards. Therefore, the EIMS/MS daughter ion spectra of the molecular ions of the Ni(II), VO(II) and metal-free CAP-6-Me (4a) and CAP-7 (6) were obtained at least four times over a two-year period. The daughter spectra of the M⁺⁺ ion of the compounds were quite reproducible. In particular, the pattern of daughter ions
Table 2-1. EIMS/MS Daughter Ion Spectra of the Molecular Ions of CAP Porphyrins.

<table>
<thead>
<tr>
<th>Compound</th>
<th>M.W.</th>
<th>% Abundance (Relative to M-15 = 100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M-15</td>
</tr>
<tr>
<td>Ni²a C5</td>
<td>532</td>
<td>100</td>
</tr>
<tr>
<td>Ni²b C5</td>
<td>532</td>
<td>100</td>
</tr>
<tr>
<td>FB²b C5</td>
<td>476</td>
<td>100</td>
</tr>
<tr>
<td>FB²c C5</td>
<td>462</td>
<td>100</td>
</tr>
<tr>
<td>Ni³ C5M</td>
<td>518</td>
<td>100</td>
</tr>
<tr>
<td>VO³ C5M</td>
<td>527</td>
<td>100</td>
</tr>
<tr>
<td>Ni⁴a C6M</td>
<td>532</td>
<td>100</td>
</tr>
<tr>
<td>VO⁴a C6M</td>
<td>541</td>
<td>100</td>
</tr>
<tr>
<td>FB⁴a C6M</td>
<td>476</td>
<td>100</td>
</tr>
<tr>
<td>Ni⁴b C6</td>
<td>532</td>
<td>100</td>
</tr>
<tr>
<td>VO⁴b C6</td>
<td>541</td>
<td>100</td>
</tr>
<tr>
<td>FB⁴b C6</td>
<td>476</td>
<td>100</td>
</tr>
<tr>
<td>FB⁵a C6</td>
<td>476</td>
<td>100</td>
</tr>
<tr>
<td>FB⁵b C6</td>
<td>462</td>
<td>100</td>
</tr>
<tr>
<td>FB⁵c C6</td>
<td>490</td>
<td>100</td>
</tr>
<tr>
<td>Ni⁶ C7</td>
<td>532</td>
<td>100</td>
</tr>
<tr>
<td>VO⁶ C7</td>
<td>541</td>
<td>100</td>
</tr>
<tr>
<td>FB⁶ C7</td>
<td>476</td>
<td>100</td>
</tr>
<tr>
<td>Ni⁷ THB</td>
<td>546</td>
<td>100</td>
</tr>
<tr>
<td>VO⁷ THB</td>
<td>555</td>
<td>100</td>
</tr>
<tr>
<td>FB⁷ THB</td>
<td>490</td>
<td>100</td>
</tr>
</tbody>
</table>

* The ions were renormalized to [M-15]⁺ = 100% to clarify the correlation. In the "M-X" columns, ions of less than 1% relative abundance are not listed. FB = Free base; C5, C6, C7 = CAP-5, CAP-6, CAP-7 porphyrins respectively; C5M, C6M = CAP-5-Me; CAP-6-Me respectively. Refer to Figure 2-1 for structures of molecules.
Figure 2-2: Selected EIMS/MS daughter ion spectra of the molecular ions of CAP porphyrin standards: a) Ni(II) $\text{C}_{32}$ \text{CAP-5} 2a (m/z 532); b) VO(II) $\text{C}_{31}$ \text{CAP-5-Me} 3 (m/z 527); c) CAP-6 5a (m/z 476); d) CAP-6-Me 4a (m/z 476); e) Ni(II) \text{CAP-7} 6 (m/z 532); f) VO(II) THB 7 (m/z 555).
in the \([\text{M-43}]^+\) to \([\text{M-45}]^+\) region of the spectrum varied less than \(\pm 10\%\) in relative abundance.

All of the daughter ion spectra show \([\text{M-15}]^+\) and \([\text{M-30}]^+\) daughter ions, which are formed by the classic \(\beta\)-cleavage of ethyl substituents of the macrocycle. CAP porphyrins with different isocyclic rings produced either different daughter ions or different distributions of daughter ions at \([\text{M-43}]^+\), \([\text{M-44}]^+\) and \([\text{M-45}]^+\). A ternary plot of the relative abundances of the three ions for the 22 compounds listed in Table 2-1 demonstrates graphically that the CAP porphyrin skeletal types can be distinguished using these three ions (Figure 2-3). For this purpose, the sum of the relative abundances of the \([\text{M-43}]^+\), \([\text{M-44}]^+\) and \([\text{M-45}]^+\) ions was set to equal 100%. The differences for the CAP-5, CAP-6, CAP-7 and THB porphyrins will now be discussed in turn.

Ni(II), VO(II) and metal-free CAP-5 porphyrins (2) all show \([\text{M-44}]^+\) as either the sole daughter ion or the dominant daughter ion in the \([\text{M-43}]^+\) to \([\text{M-45}]^+\) region of the spectrum. The \([\text{M-45}]^+\) ion may also occur as a minor daughter ion, with the highest relative abundance occurring for the Ni(II) compounds (Table 2-1; Figure 2-2). All of the five CAP-5 porphyrins studied clustered together in the ternary diagram (Figure 2-3). Varying the substitution pattern, carbon number and the nature of the chelated metal ion (if any) of the CAP-5 porphyrins produced only minor variations in the daughter ion spectra of the \(\text{M}^+\) ion. Nevertheless, the \([\text{M-44}]^+\) was always the dominant ion, and the \([\text{M-43}]^+\) ion was
Figure 2-3: Ternary diagram of the relative intensities of the (M-43)$^+$, (M-44)$^+$ and (M-45)$^+$ daughter ions of the $M^+$ ions of Ni(II) VO(II) and metal-free CAP porphyrin standards listed in table 2-1.
not detected (Table 2-1). Furthermore, the \([M-44]^+\) ion was always more abundant than the \([M-30]^+\) ion.

Only one CAP-5 porphyrin with a methyl substituent on the isocyclic ring was studied (3). The daughter ion spectra of the \(M^+\) ion of both the Ni(II) and VO(II) complexes of 3 revealed \([M-44]^+\) as the dominant ion in the \([M-43]^+\) to \([M-45]^+\) region of the spectra (Figure 2-2, Table 2-1). These compounds clustered with the CAP-5 porphyrins in the ternary plot (Figure 2-3).

The Ni(II), VO(II) and metal-free CAP-6 porphyrins 4b, 5a, 5b and 5c show \([M-45]^+\) as the dominant daughter ion in this region of the EIMS/MS daughter ion spectrum. In contrast to the CAP-5 porphyrins, the fragmentation pattern can vary more significantly with the nature of the chelated metal ion, the substitution pattern, and the carbon number. The principal variations in these daughter ion spectra lie in the relative abundances of the \([M-45]^+\) and \([M-30]^+\) ions and the presence or absence of \([M-44]^+\) as a minor daughter ion (Table 2-1). For example, the daughter ion spectrum of the metal-free porphyrin 5a shows \([M-30]^+\) to be more abundant than \([M-45]^+\), which is the reverse of what is observed for both the Ni(II) and VO(II) complexes (Table 2-1). Despite these variations, the CAP-6 porphyrins clustered together quite well on the ternary plot (Figure 2-3).

The presence of a methyl substituent on the 6-membered isocyclic ring produces a significantly different fragmentation pattern in the \([M-43]^+\) to \([M-45]^+\) region of the spectrum. For Ni(II), VO(II) and metal-free 4b, the \([M-44]^+\) and \([M-45]^+\) daughter ions occur in similar relative abundances (Table 2-1; Figure
The three compounds formed a cluster between those of the CAP-5 and CAP-6 porphyrins (Figure 2-3).

Only one CAP-7 porphyrin was studied (6). For the Ni(II), VO(II) and metal-free forms the daughter ions [M-43]+, [M-44]+ and [M-45]+ were all observed in relative abundance order of [M-43]+ > [M-45]+ ≥ [M-44]+. The ions in this region were less abundant than the [M-30]+ daughter ion for the metal complexes, but were of similar abundance for the metal-free complexes (Table 2-1; Figure 2-2). The three porphyrins formed a cluster that was well separated from those of the other porphyrin skeletal types (Figure 2-3).

The Ni(II), VO(II) and metal-free forms of the THB porphyrin (Z) all show [M-43]+ as the dominant ion in this part of the EIMS/MS daughter ion spectrum. For the metal-free porphyrin, this daughter ion is of lower relative abundance than the [M-30]+ ion, which is the reverse of what is observed for the metal complexes of Z (Table 2-1; Figure 2-2). The three THB porphyrins formed a cluster that was quite distinct from those of the other CAP porphyrins.

The THB porphyrins were similar to the CAP-7 porphyrins in that the [M-43]+ ion appeared as the most intense ion of the [M-43]+, [M-44]+, and [M-45]+ cluster. In THB, there is a six-membered cycloalkyl group on the pyrrole ring; the CAP-7 porphyrin has the ring attached at the meso and β carbons of the porphyrin (Figure 2-1). In both of these cases, the dominant [M-43]+ ion could correspond to the loss of C₃H₇. The [M-43]+ could be the result result of
\(\alpha\)-cleavage at one end of the exocyclic ring and \(\beta\)-cleavage at the other with the transfer of a single hydrogen.

In addition to the daughter ions already discussed, two other daughter ions are often observed, \([M-31]^+\) and \([M-59]^+\). The \([M-31]^+\) ion was detected in many of the spectra. Typically, it occurred in higher relative abundance in the spectra of the CAP-6 porphyrin series than in the other skeletal types; however, it was most abundant in the spectrum of the \(C_{31}\), CAP, 2c. At this point it does not appear likely that this daughter ion will be valuable for distinguishing the different types of CAP porphyrins because there is no obvious relationship between the porphyrin skeletal types and its relative abundance. The \([M-59]^+\) ion is of rather more interest. It was observed in the daughter ion spectra of the molecular ions of all the porphyrins studied except for the metal complexes of the THB porphyrin, 7. It was of similar relative abundance to the \([M-43]^+\) ion for the CAP-7 porphyrins (Table 2-1). It was of lower relative abundance for the CAP-5 porphyrins except for the metal-free form of 2b. There was no clear pattern between the relative abundances of the \([M-59]^+\) and the \([M-45]^+\) ions for CAP-6 porphyrins.

In summary, the data indicate that careful analysis of EIMS/MS daughter ion spectra of molecular ions permits the assignment of the isocyclic ring. The size of the isocyclic ring may be ascertained using the following diagnostic features:
(a) CAP-5, 2, and CAP-5-Me porphyrins, 3, display a dominant [M-44]$^+$ ion in 
(b) CAP-5 and CAP-5-Me porphyrins are not distinguished by EIMS/MS.
(c) CAP-6 porphyrins, 5, are characterized by a dominant [M-45]$^+$ in the this 
region of the spectrum.
(d) The presence of [M-45]$^+$ and [M-44]$^+$ ions in similar relative abundance is 
indicative of CAP-6-Me porphyrins, 4.
(e) The CAP-7 porphyrins show the following pattern of daughter ions: [M-43]$^+$ 
$> [M-45]^+ > [M-44]^+$.
(f) THB porphyrins are characterized by a dominant [M-43]$^+$ ion.

Comparison With the Daughter Ion Spectra of Molecular Ions of Geoporphyrins

It is appropriate to apply the results from the present study to the analysis 
of geoporphyrin mixtures. Thus, the daughter ion spectra of the M$^+$ ions of nine 
geoporphyrins isolated from a total organic extract (bitumen-I) of the New Albany 
Shale were re-examined [Beato et al., 1991]. The relative intensities of the 
[M-43]$^+$, [M-44]$^+$ and [M-45]$^+$ daughter ions are summarized in Table 2-2 together 
with the re-normalized intensities, which were used for a ternary plot of the three 
ions (Figure 2-4) Daughter ion spectra of the Ni(II) C$_{30}$, Ni(II) C$_{33}$, Ni(II) C$_{34}$ and 
VO(II) C$_{32}$ CAP porphyrins together with two metal-free C$_{30}$ CAP porphyrins are 
shown in Figure 2-5. The selection criteria and the salient features of the 
EIMS/MS analysis of the CAP geoporphyrins are discussed below.

The daughter ion spectra of the M$^+$ ions of Ni(II) and VO(II) C$_{30}$ and C$_{32}$ 
CAP porphyrins were selected because the spectra are typical of those observed 
for CAP geoporphyrins. The spectra of the demetallated C$_{30}$ CAP porphyrins were 
studied because they are isomeric, and therefore might be of different skeletal

<table>
<thead>
<tr>
<th>Compound</th>
<th>M.W.</th>
<th>% Relative Abundance*</th>
<th></th>
<th>Re-normalized**</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M-43</td>
<td>M-44</td>
<td>M-45</td>
<td>M-43</td>
</tr>
<tr>
<td>NiC₃₀</td>
<td>504</td>
<td>1</td>
<td>49</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>NiC₃₂</td>
<td>532</td>
<td>4</td>
<td>47</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>NiC₃₃</td>
<td>546</td>
<td>18</td>
<td>10</td>
<td>4</td>
<td>56</td>
</tr>
<tr>
<td>NiC₃₄</td>
<td>560</td>
<td>18</td>
<td>-</td>
<td>1</td>
<td>95</td>
</tr>
<tr>
<td>VOC₃₀</td>
<td>513</td>
<td>-</td>
<td>46</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>VOC₃₂</td>
<td>541</td>
<td>-</td>
<td>42</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>VOC₃₃</td>
<td>555</td>
<td>2</td>
<td>39</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>FBC₃₀L§</td>
<td>448</td>
<td>10</td>
<td>15</td>
<td>50</td>
<td>13</td>
</tr>
<tr>
<td>FBC₃₀M§</td>
<td>448</td>
<td>1</td>
<td>22</td>
<td>13</td>
<td>3</td>
</tr>
</tbody>
</table>

* The ions were renormalized to [M-15]⁺ = 100% for comparison with the data in Table 2-1.
** Re-normalized so that Σ[M-43]⁺ + [M-44]⁺ + [M-45]⁺ = 100% for ternary plot (Figure 2-4)
§These isomeric CAP porphyrins were obtained by demetallation of the total Ni(IIlnn) porphyrin mixture followed by thin layer chromatography.
FB = Free base; L = Less polar isomer; M = More polar isomer.
Figure 2-4: Ternary diagram of the relative intensities of the (M-43)$^+$, (M-44)$^+$ and (M-45)$^+$ daughter ions of the M$^+$ ions of VO(II) $C_{30}$, $C_{32}$ and $C_{33}$ CAP geoporphyrins, and Ni(II) $C_{30}$, $C_{32}$, $C_{33}$ $C_{34}$ CAP geoporphyrins together with two metal-free $C_{30}$ CAP geoporphyrin isomers from New Albany bitumen-I.
Figure 2-5: Selected EIMS/MS daughter ion spectra of the molecular ions of CAP geoporphyrins: 
a) Ni(II) C\textsubscript{30} CAP (m/z 504); b) VO(II) C\textsubscript{32} CAP-5-Me (m/z 541); c) Ni(II) C\textsubscript{33} CAP (m/z 546); d) Ni(II) C\textsubscript{34} CAP (m/z 560); e) Less polar metal free C\textsubscript{30} CAP isomer (m/z 448); f) more polar C\textsubscript{30} CAP isomer (m/z 448) from New Albany bitumen-I.
types. The spectra of the Ni(II) and VO(II) C₃₃ CAP porphyrins were very different, and clearly warranted re-investigation. The Ni(II) C₃₄ CAP porphyrin was analyzed because it is an example of a high carbon-number geoporphyrin porphyrin [Johnson et al., 1986], which could not have been readily derived from chlorophyll a via a Treibs' type degradation pathway [Treibs, 1936].

The VO(II) and Ni(II) C₃₀ and C₃₂ CAP porphyrins displayed abundant [M-44]⁺ ions (Figure 2-5, Table 2-2), indicating that were CAP-5 or possibly CAP-5-Me porphyrins. On the ternary diagram (Figure 2-4), they formed a cluster in the same place as the CAP-5 and CAP-5-Me standards. This is not surprising because it has long been assumed that the CAP-5 porphyrins (DPEP) are usually the dominant skeletal type of CAP porphyrins.

The presence of the dominant [M-44]⁺ ion in the more polar metal free C₃₀ CAP isomer also was indicative of a CAP-5 porphyrin. In contrast, the [M-45]⁺ daughter ion was dominant in the less polar isomer, which indicates that it is a CAP-6 compound (Figure 2-5, Table 2-2). On the ternary diagram, the less polar isomer lay in the domain of the CAP-6 porphyrins. The more polar isomer lay close to the CAP-5 porphyrins (Figure 2-4).

The daughter ion spectra of the Ni(II) and VO(II) C₃₃ porphyrins were quite different from each other. The Ni(II) complex showed [M-43]⁺ as the most intense of these ions (Figure 2-5). On the ternary diagram it lay in the domain of the CAP-7 porphyrins (Figure 2-4). In contrast, the VO(II) complex lay in the domain of the CAP-5 porphyrins (Figure 2-4).
The daughter ion spectrum of the Ni(II) $C_{34}$ complex is less easily interpreted (Figure 2-5, Table 2-2). The spectrum shows an intense $[M-43]^+$ daughter ion, and the compound lay within the domain of the THB porphyrins in the ternary plot (Figure 2-4) These data could conceivably indicate that the porphyrin is a CAP-7 or THB type species; however, such speculation must be treated with caution. The $[M-43]^+$ might also be the product of $\beta$-cleavage of a butyl or isobutyl moiety, which would sharply skew the fragmentation pattern. Nevertheless, it is intriguing to note that in the daughter ion spectra of the molecular ions of all higher carbon number ($\geq C_{34}$) porphyrins studied from both New Albany Shale and Boscan oil the $[M-44]^+$ and $[M-45]^+$ daughter ions were always of very low relative abundance (Johnson et al., 1986; Quirke et al., 1989).

The above data demonstrate the potential value of tandem mass spectrometry in the characterization of CAP geoporphyrins. It is essential to perform chromatographic separations to avoid the problem of interpreting the daughter ion spectra of unresolved isomers. The use of the ternary diagrams may provide the researcher with a means of determining whether such a problem has arisen. Clearly, on-line high performance liquid chromatography/tandem mass spectrometer (LC/MS/MS) would be an ideal way to effect the analysis of geoporphyrin mixtures.
Conclusions

The EIMS/MS daughter ion spectra of the molecular ions of CAP porphyrins have been obtained. The [M-43]^+, [M-44]^+ and [M-45]^+ daughter ions are the most valuable ions for distinguishing the skeletal type of CAP porphyrins. An intense [M-44]^+ daughter ion relative to [M-43]^+ or [M-45]^+ is indicative of a CAP-5 or CAP-5-Me porphyrin. CAP-6 porphyrins display [M-45]^+ as the most intense of these daughter ions. CAP-6-Me porphyrins show [M-44]^+ and [M-45]^+ daughter ions in similar relative abundances. All three ions are present in the daughter ion spectra of CAP-7 porphyrins with the [M-43]^+ ion being the most intense. The THB porphyrins are characterized by an intense [M-43]^+ daughter ion, the other two ions being present in very low relative abundance. The need to separate the isomeric CAP porphyrins to avoid problems in interpretation of the daughter ion spectra of the M^+ ions would make LC/MS/MS the ideal method for the analysis of geoporphyrin mixtures.

In normalizing the intensities of the [M-43]^+, [M-44]^+, and [M-45]^+ daughter ions and plotting them in a ternary diagram, distinct clusters of data points correspond to the various CAP skeletal types. When applied to data already obtained for partially-separated mixtures of geoporphyrins, the technique showed promise for the analysis of geoporphyrins. Most of the geoporphyrin data plotted in the CAP-5/CAP-5-Me region of the ternary diagram. However, two of the data points for high-carbon number porphyrins (the Ni(II) C_{33} and the Ni(II) C_{34}) plotted
in the CAP-7 and THB regions, respectively, of the ternary diagram. This indicates CAP-7 and THB porphyrins may be present in Shales and oil deposits (in opposition to the Treibs hypothesis); previous studies in our group [Johnson et al., 1986] were unable to confidently propose the structures of the high-carbon number porphyrins since standard compounds were not available at the time.
CHAPTER 3
ELECTRON IONIZATION TANDEM MASS SPECTROMETRY OF 5-NO₂ OCTAETHYLPORPHYRIN: UNUSUAL FRAGMENTATIONS AND THE INFLUENCE OF CHELATED DIVALENT METAL IONS

Introduction

Porphyrins are of considerable importance in many areas of chemistry and biology. For example, heme, the prosthetic group in hemoglobin and most cytochromes, plays a crucial role in biological transportation of oxygen and electrons respectively [Marks, 1969; Lemberg and Barret, 1973]. Perhaps the most important property of porphyrin macrocycles is their ability to chelate with virtually all the metallic elements [Buchler, 1975].

Spectroscopic Studies of Nitroporphyrins

It has long been known that the nature of the chelated metal ion, the peripheral substituents, and the axial ligand can potentially alter the spectroscopic, redox, and conformational properties of porphyrins. These properties are of foremost importance in determining the activity of the porphyrins in biological reaction centers. In particular, much research has been performed
recently to determine the importance of the non-planar conformation of many biologically active porphyrins [Trunrud et al., 1986; Horning et al., 1986; Barkigia et al., 1988; Deisenhofer, 1989]. An understanding of the structure-activity relationship of distorted biological porphyrins aids in the development of biomimetic photocatalysts [Shelnutt and Trudell, 1989].

Recent studies on meso-nitro-octaalkyl porphyrins have indicated that the compounds are of considerable value as biological models. Medforth et al. [1990; 1992] characterized sterically strained non-planar porphyrins containing bulky peripheral substituents. However, the non-planarity of meso-nitroporphyrins has brought them interest in the investigation of properties of non-planar porphyrins. This is especially true since the synthesis of nitroporphyrins is considerably easier than that for other non-planar systems and electron withdrawing groups can facilitate axial ligation [Senge, 1993]. Anderson et al. [1993] proposed that the small perturbation of the mono-nitro porphyrins may more accurately represent the distortion present in the protein-bound environment than the more highly substituted porphyrins. One prevalent example is the study of cytochrome c [Shelnutt et al., 1992; Anderson et al., 1993; Hobbs et al., 1994]. Extensive spectroscopic studies have been performed on 5-nitro octaethylporphyrin (5-NO₂ OEP) and its divalent metal complexes [Gong and Dolphin, 1984; Stanley, 1990; Wu et al., 1991; Liu, 1993]. The effect of the coordinated metal on the porphyrin macrocycle has been studied by various spectroscopic techniques. Infrared (IR) [Ogoshi and Yoshida, 1971; Ogoshi et al.,
Mass Spectrometry of Nitroporphyrins

Porphyrrins have been of interest in mass spectrometry as they have proved useful in investigating a variety of phenomena. The high relative abundance of doubly charged ions occurring in electron ionization mass spectrometry (EIMS) of porphyrins and their fragmentation pathways were studied by Beato et al. [1989a]. In chemical ionization mass spectrometry (CIMS), surface-induced decomposition of porphyrins into mono-, di-, and tripyrrolic units has been observed [Shaw et al., 1981; Beato et al., 1989b; Van Berkel et al., 1989a,b]. Recently, Van Berkel et al. [1992] reported that Ni(II) porphyrins undergo electrochemistry in the electrospray needle in electrospray ionization mass spectrometry (ESMS). Yan et al. have studied the effect of the meso substituent on the linear dependence of the appearance of [M+H]^+ ions (basicity) in fast-atom bombardment (FAB) mass spectrometry of OEPs [1991]. The effect of metal ions on ammonia CI mass spectra of metalloporphyrins has been
investigated [Beato et al., 1989b] In all of the above studies, the characteristics of the chelated metal ion, if any, modified the mass spectra.

There have been only limited reports on the EIMS analysis of meso (bridge) substituted porphyrins. Jackson et al. presented the first detailed paper on the mass spectrometry of porphyrins in 1965. Budzikiewicz [1978] performed a few studies on the mass spectrometry of meso-substituted porphyrins. In general it was believed that the presence of a meso substituent did not modify the fragmentation pathway significantly. However, the studies presented in Chapter 2 of this dissertation on cyclo-alkano-porphyrins (CAP) indicate that unusual fragmentations can occur. Clezy et al. determined that the presence of a nitro group at one of the bridge carbons modified the EIMS spectrum substantially because the normally encountered δ-cleavage of pyrrolic alkyl substituents was suppressed [1974]. The fragmentations may be a result of the meso substituent or of the conformation of the isocyclic ring as a whole [Jackson et al., 1965; Clezy et al., 1974; Smith, 1975; Budzikiewicz 1978]. Investigation of such novel fragmentations by electron ionization tandem mass spectrometry (EIMS/MS) is of fundamental interest since it may be possible to relate the mass spectrometric data to other characteristic spectroscopic properties.

The mass spectra of metalloporphyrins are markedly influenced by the nature of the chelated metal ion. Beato et al. [1989b] showed that the relative abundance of doubly charged ions to singly-charged ions in the El mass spectra of metallated complexes of octaethylporphyrin OEP is related to the chelated
metal (as is discussed in Chapter 1 of this thesis). The singly charged region of
the mass spectrum, however, remains relatively unaffected by the identity of the
chelated metal ion. Undoubtedly, the role of the metal ion in the mass spectra
of metalloporphyrins warrants extensive investigation.

The present study is limited to an investigation of the divalent metal complexes of porphyrins for two reasons. First, trivalent or tetravalent complexes
bear axial ligands that are often exchangeable. Thus, the homogeneity of such
complexes may be hard to guarantee. Second, it is known that higher valent
metal complexes may be transformed into divalent metal complexes during
volatilization [Edwards et al., 1970].

In this Chapter are presented the results of EI MS, EI MS/MS and electron
ionization high resolution mass spectrometric (EI HRMS) analysis of 5-NO₂
octaethylporphyrin. A novel fragmentation involving the scission of the porphyrin
macrocycle is presented and the implications are discussed. In contrast to the
mass spectra for OEP and its metal complexes, the insertion of metals into 5-NO₂-
OEP was found to have a strong influence on the types of fragmentation
encountered in the singly charged region of the mass spectrum. The divalent
metal complexes studied include Co, Cu, Zn, Ni, Pd, Ag and Mg. The effect of
the metals on the fragmentations is to be discussed in detail; the relation of the
mass spectral data to other metal-dependent spectroscopic data is also
examined.
Compound Preparation

**Synthesis of NO$_2$ porphyrins.** 5-NO$_2$ octaethylporphyrin was prepared by the method of Bonnett and Stephenson [1965]. The method involves the electrophilic substitution of the nitro group at one of the bridge carbons using fuming nitric and acetic acid as the nitrating agent. A mixture of 5-NO$_2$ OEP and the 5,10 and the 5,15 di-NO$_2$ isomers of OEP were obtained.

The 5-NO$_2$ was purified by column chromatography. Using the slurry method, a ten inch bed of silica (200-400 mesh) was packed in 4:1 hexane/toluene. The compounds were then eluted with 4:1 hexane/toluene, increasing the polarity as needed. The first fraction, which eluted with 3:1 hexane/toluene, contained the di-NO$_2$ OEP isomers. As the polarity increased, the 5-NO$_2$ OEP eluted in 2:1 hexane/toluene in a distinct band. The remaining porphyrin band, which eluted with 1:1 hexane/toluene, contained the unreacted OEP. The column was finally washed with dichloromethane and methanol. The second fraction, which contained the pure 5-NO$_2$ OEP, was evaporated under vacuum and recrystallized from 1:1 dichloromethane/methanol.

**Synthesis of isotopically labelled porphyrins.** Octaethylporphyrin labelled with $^{15}$N at all four porphyrinic nitrogens was prepared by the method of Callot et al. using $^{15}$N ammonium chloride (Cambridge Isotopic Laboratories) as the initial
source of nitrogen [1983]. Octaethylporphyrin labelled with $^{13}$C at each of the four bridge positions was also prepared by the method of Callot [1983]. Each of these two isotopically labelled compounds were nitrated by the method of Bonnett and Stephenson [1965] and purified and recrystallized as above.

The structure of 5-NO$_2$ OEP (or meso nitro OEP) is shown in Figure 3-1. The sites of isotopic labelling with $^{13}$C and $^{15}$N have been marked with symbols for clarification.

**Synthesis of Ni(II) and Co(II) 5-NO$_2$ OEP.** The 5-NO$_2$ OEP (50 mg) and nickel (II) acetate (100 mg) were dissolved in dimethyl formamide (15 mL). The solution was refluxed and was monitored by UV-visible spectrophotometry. When the reaction was completed (after about one hour), the solution was diluted with water and extracted with dichloromethane. The organic layer was separated and washed with water. The organic layer was then evaporated under vacuum and the porphyrin was recrystallized as described above. TLC analysis for Ni(II) 5-NO$_2$ OEP were carried out using hexane/toluene (1:1;v:v) as the mobile phase.

Co(II) 5-NO$_2$ OEP was prepared by the same method using cobalt (II) acetate. The solvent system had to be modified to dissolve the porphyrin and salt by using 2 mL of dichloromethane and 10 mL of methanol. TLC was carried out using hexane/toluene (1:1;v:v) as the mobile phase.

**Synthesis of Zn(II) and Cu(II) 5-NO$_2$ OEP.** The 5-NO$_2$ OEP (50 mg) was dissolved in dichloromethane (25 ml) and a solution of zinc acetate (100 mg) in methanol was added. The mixture was refluxed for one hour and the product was
Figure 3-1: Structure of 5-NO₂ OEP. Sites of $^{13}$C labelling indicated by (○). Sites of $^{15}$N labelling indicated by (★).
isolated by distilling off the dichloromethane and adding methanol. The Zn(II) 5-NO$_2$ OEP was filtered off and recrystallized as described above. The purity of the Zn(II) 5-NO$_2$ OEP was determined by TLC with hexane/toluene (1:1;v:v) as the mobile phase.

Cu(II) 5-NO$_2$ OEP was synthesized as above using copper (II) acetate. TLC analyses were carried out using hexane/toluene (1:1; v:v).

**Synthesis of Mg 5-NO$_2$ OEP.** The 5-NO$_2$ OEP (50 mg) and magnesium perchlorate (200 mg) were dissolved in pyridine. The mixture was refluxed for two hours and the reaction was monitored by UV-visible absorption until complete. Upon completion the pyridine was removed under vacuum and the product was washed several times with water. The product was then dissolved in dichloromethane which had been eluted through a grade V alumina column to remove any traces of acid. The Mg(II) 5-NO$_2$ OEP was then recrystallized in dichloromethane/methanol (1:1;v:v) which had also been filtered through alumina to remove acid. TLC analysis for Mg(II) 5-NO$_2$ OEP was carried out using 10% ethyl acetate in dichloromethane as the mobile phase.

**Synthesis of Ag(II) 5-NO$_2$ OEP.** The mixture of 5-NO$_2$ OEP (30 mg) and AgNO$_3$ (100 mg) was dissolved in DMF (7 mL). The solution was refluxed under an argon atmosphere and monitored by UV-visible spectroscopy. After the reaction was completed (ca. 15 min), the mixture was diluted with 100mL of water and extracted with methylene chloride. The porphyrin went into the organic phase. Any solid residue (primarily inorganic byproducts) was filtered off. The solvent
was evaporated and the residue was recrystallized with dichloromethane/methanol (1:1;v:v).

Synthesis of Pd(II) OEP. 5-NO₂ OEP (50 mg) and PdCl₂ (50 mg) were dissolved in benzonitrile (7 mL), and the mixture was refluxed under argon atmosphere until metallation was completed as determined by UV-visible spectra (ca. 15 min). The mixture was chilled and vacuum filtered. The solid product was recrystallized from pure toluene.

Methods

All mass spectra presented here were obtained with a Finnigan MAT TSQ 45 triple quadrupole mass spectrometer. Porphyrin samples were dissolved in dichloromethane and 1 μL aliquots of the solution were deposited onto a direct exposure probe (DEP) filament. The DEP was heated from 100 to 600 °C at 600 °C min⁻¹. The rapid rate of heating was found to reduce the amount of thermal elimination of NO₂ observed in EIMS.

The spectra were obtained under standard EI conditions. The filament emission current was maintained at 0.3 mA and the electron energy was 70 eV for all experiments. For the initial EIMS experiments quadrupole 1 (Q1) was scanned from m/z 50 to 800 in 0.8 s with Q2 and Q3 passing all masses. Tuning for MS/MS was achieved with Co(II) OEP in a capped aluminum solids probe vial
heated to 275°C. The MS/MS experiments were performed with a collision gas pressure of 1.6 mTorr (argon) and a collision energy of 24.8 eV.

A Finnigan MAT 95 high resolution mass spectrometer was used for determination of exact masses for El fragment ions. The MAT 95 was tuned for a resolution of 7000. In this case, the samples were desorbed from a water-cooled solids probe. The high resolution data were used as an aid in determining the composition of certain fragment ions.

Results and Discussion

The singly charged fragment ions of OEP arise primarily from a sequence of successive β-cleavages of the peripheral ethyl groups (Figure 3-2a). This fragmentation pathway was confirmed by analysis of the metastable ions by Clezy et al. [1974]. The porphyrin macrocycle remains intact; doubly charged ions are present in high relative abundance as would be expected. The work presented here concentrates only on the singly charged region of the mass spectrum of 5-NO₂ OEP.

Analysis of Free-base 5-NO₂ OEP

With a nitro group attached to the bridge carbon position, the fragmentation becomes significantly altered. In our MS and MS/MS studies of
Figure 3-2: El mass spectra of 5-NO$_2$ OEP: a) MS of OEP; b) MS of 5-NO$_2$ OEP; c) Daughter spectrum of the molecular ion of 5-NO$_2$ OEP; d) Parent spectrum of the m/z 375 ion of 5-NO$_2$ OEP.
free-base 5-NO$_2$-OEP, both normal El and daughter ion spectra (Figures 3-2b and 3-2c) showed a significant [M-17]$^+$ peak owing to loss of OH$^-$ as was reported by Clezy et al. [1974] in their EIMS work. The mechanism that has been proposed by Clezy et al. for the loss of OH$^-$ is shown in Figure 3-3. The elimination of a hydrogen from the $\beta$-substituent adjacent to the nitro group occurs with a seven-membered cyclic rearrangement, giving an ion at m/z 562.

From the El mass spectrum of 5-NO$_2$-OEP (Figure 3-2b) we see that peaks corresponding to $\beta$-cleavage ions (e.g. m/z 564) are minor, whereas loss of OH$^-$ (m/z 562) as well as $\alpha$-cleavage (m/z 550) result in major peaks. A peak at m/z 533 has been attributed to loss of NO$_2$.

We have found evidence for a novel ring fragmentation of 5-NO$_2$-OEP. Our EIMS analysis showed a significant peak at m/z 375 which we have attributed to loss of a pyrrole unit. Such fragmentation processes for substituted porphyrins have previously been observed only as a surface-induced fragmentation process in CIMS [Beato et al., 1989a] (refer to Chapter 1 of this dissertation for the detailed discussion of this phenomenon).

The EIMS analysis of 5-NO$_2$ OEP is not easy to perform because the nitro group may be thermally eliminated in the solids probe vial or ion source during the heating required for volatilization, with intermolecular hydrogen abstraction forming OEP. This problem can be avoided if it is possible to heat the sample rapidly. Therefore the use of a direct exposure probe is to be recommended as opposed to use of the solids probe. The EIMS of 5-NO$_2$ OEP is shown in Figure
Figure 3-3: Loss of OH radical from 5-NO2 OEP as reported by Clezy et al. [1974].
3-2b. The spectrum is broadly similar to that reported by Clezy et al. in 1974. In the present study, there was little contamination from thermal decomposition of the 5-NO$_2$ OEP, as evidenced by the very low abundance of the M$^+$ for OEP (m/z 534).

The singly charged region of the molecule is complex, and is substantially different from that of OEP. The presence of the m/z 564 ion as a minor component in both the EIMS and the EIMS/MS daughter ion spectrum of the M$^+$ ion of 5-NO$_2$ OEP confirmed that the $\beta$-cleavage fragmentation pathway is suppressed in this molecule. The eight most abundant fragment ions observed were m/z 562, 550, 533, 523, 522, 504, 489, and 375. It was not obvious which of these ions were produced directly from the molecular ion, M$^{+\cdot}$ (m/z 579), or from thermal decomposition occurring in the sample vial or in the ion source. Therefore the EIMS/MS daughter ion spectrum of the M$^{+\cdot}$ ion of 5-NO$_2$ OEP was obtained. The spectrum is shown in Figure 3-2c; the most abundant daughter ions were m/z 564, 562, 550, 535, 533, 523, 522, 521, and 375. The origin of each of these ions will be discussed in turn.

The [M-17]$^+$ daughter ion, m/z 562, is formed by loss of a hydroxy radical via a rearrangement such as that proposed by Clezy et al. [1974] and is discussed above (Figures 3-2c and 3-3). The ions m/z 550, [M-29]$^+$, and 533, [M-46]$^+$, are generated by $\alpha$-cleavage of ethyl and nitro groups, respectively, and are analogous to the same ions produced in the EIMS spectrum.
The origins of the daughter ions m/z 535, 523, 522, 521 and 375 are more difficult to assign. In fact, it was essential to obtain electron ionization high resolution mass spectra (EIHRMS) to determine the elemental composition of the neutrals lost and the daughter ions. Table 3-1 shows the exact mass, \( \Delta \) (the difference in millimass units (mmu) from the actual mass of the ion reported), percent relative abundance, and the ion composition for the ions of interest in the mass spectra of 5-NO\(_2\) OEP and its isotopically labelled analogs. For the \(^{13}\text{C}_4\) and \(^{15}\text{N}_4\) labelled compounds, only the ions pertinent to the ring cleavage process are shown. The EIHRMS analyses confirmed the interpretation of the m/z 562, 550 and 533 daughter ions. Unfortunately, although the m/z 535, \((\text{M}-44)^+\) ion is quite abundant in the EIMS/MS spectrum (Figure 3-2b), it is of very low relative abundance in both the EIMS and the EIHRMS spectrum; therefore, it was not possible to determine either the elemental composition or the pathway to this ion. The m/z 522 daughter ion is the result of a loss of \(\text{C}_3\text{H}_6\text{O}\) from the \(\text{M}^+\) ion. To account for this ion, it is necessary to invoke a cleavage of a carbon from the macrocycle, as is shown in Figure 3-4. The mechanism of this fragmentation will be discussed in more detail below, as it appears that the m/z 522 ion is an important intermediate in the ring scission process (see Figure 3-5) and the formation of the m/z 375 ion. A similar pathway may be used to explain the formation of the ion m/z 521 \((\text{M}-58)^+\), in which the neutral lost has the elemental composition \(\text{C}_3\text{H}_6\text{O}\). The m/z 523 ion is the \(^{13}\text{C}\) peak for the ion at m/z 522. The relative intensities of the two ions confirm this. The most intriguing daughter ion
<table>
<thead>
<tr>
<th>Exact Mass (Ammu)</th>
<th>% Relative Abundance</th>
<th>Loss Composition</th>
<th>Ion Composition</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>579.3600 (2.7)</td>
<td>100.00</td>
<td>-</td>
<td>C\textsubscript{30}H\textsubscript{45}O\textsubscript{2}N\textsubscript{5}</td>
<td>molecular ion</td>
</tr>
<tr>
<td>564.3384 (4.6)</td>
<td>1.94</td>
<td>CH\textsubscript{3}</td>
<td>C\textsubscript{35}H\textsubscript{42}O\textsubscript{2}N\textsubscript{3}</td>
<td>(\beta)-cleavage</td>
</tr>
<tr>
<td>562.3505 (-4.1)</td>
<td>6.34</td>
<td>OH</td>
<td>C\textsubscript{36}H\textsubscript{46}ON\textsubscript{5}</td>
<td>Clezy et al., 1974</td>
</tr>
<tr>
<td>unlabelled</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>550.3270 (8.8)</td>
<td>4.47</td>
<td>C\textsubscript{2}H\textsubscript{5}</td>
<td>C\textsubscript{34}H\textsubscript{46}O\textsubscript{2}N\textsubscript{5}</td>
<td>(\alpha)-cleavage</td>
</tr>
<tr>
<td>549.3505 (-12.7)</td>
<td>4.45</td>
<td>CH\textsubscript{2}O</td>
<td>C\textsubscript{35}H\textsubscript{43}ON\textsubscript{5}</td>
<td>-</td>
</tr>
<tr>
<td>compound</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>523.3250 (-0.8)</td>
<td>1.82</td>
<td>C\textsubscript{3}H\textsubscript{6}O</td>
<td>C\textsubscript{32}H\textsubscript{40}^{13}CON\textsubscript{5}</td>
<td>(^{13})C peak</td>
</tr>
<tr>
<td>522.3218 (-0.5)</td>
<td>4.08</td>
<td>C\textsubscript{3}H\textsubscript{6}O</td>
<td>C\textsubscript{33}H\textsubscript{40}ON\textsubscript{5}</td>
<td>oxygen migration</td>
</tr>
<tr>
<td>521.3183 (2.8)</td>
<td>1.24</td>
<td>C\textsubscript{3}H\textsubscript{5}O</td>
<td>C\textsubscript{33}H\textsubscript{40}ON\textsubscript{5}</td>
<td>-</td>
</tr>
<tr>
<td>508.3131 (-7.1)</td>
<td>0.71</td>
<td>C\textsubscript{3}H\textsubscript{5}NO</td>
<td>C\textsubscript{33}H\textsubscript{40}ON\textsubscript{5}</td>
<td>-</td>
</tr>
<tr>
<td>375.2282 (-2.9)</td>
<td>0.37</td>
<td>C\textsubscript{12}H\textsubscript{16}ON\textsubscript{2}</td>
<td>C\textsubscript{24}H\textsubscript{29}O\textsubscript{3}N</td>
<td>pyrrole loss</td>
</tr>
</tbody>
</table>

\[^{13}\text{C}_4\] labelled

<table>
<thead>
<tr>
<th>Exact Mass (Ammu)</th>
<th>% Relative Abundance</th>
<th>Loss Composition</th>
<th>Ion Composition</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>538.3716 (-1.1)</td>
<td>100.00</td>
<td>-</td>
<td>C\textsubscript{32}H\textsubscript{45}O\textsubscript{2}^{13}\text{C}_4\textsubscript{4}N\textsubscript{5}</td>
<td>molecular ion</td>
</tr>
<tr>
<td>526.3344 (2.3)</td>
<td>5.19</td>
<td>C\textsubscript{3}H\textsubscript{5}O</td>
<td>C\textsubscript{29}H\textsubscript{40}O^{13}\text{C}_4\textsubscript{4}N\textsubscript{5}</td>
<td>oxygen migration</td>
</tr>
<tr>
<td>378.2400 (1.1)</td>
<td>0.92</td>
<td>C\textsubscript{11}H\textsubscript{16}^{13}\text{CON}_{2} \text{</td>
<td>C}_{21}H\textsubscript{20}O^{13}\text{C}<em>3\textsubscript{3}N</em>{3}</td>
<td>meso (^{13})C loss</td>
</tr>
</tbody>
</table>

\[^{15}\text{N}_4\] labelled

<table>
<thead>
<tr>
<th>Exact Mass (Ammu)</th>
<th>% Relative Abundance</th>
<th>Loss Composition</th>
<th>Ion Composition</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>583.3412 (4.3)</td>
<td>100.00</td>
<td>-</td>
<td>C\textsubscript{36}H\textsubscript{46}NO\textsubscript{2}^{15}\text{N}_4</td>
<td>molecular ion</td>
</tr>
<tr>
<td>526.3117 (-0.3)</td>
<td>5.04</td>
<td>C\textsubscript{3}H\textsubscript{5}O</td>
<td>C\textsubscript{33}H\textsubscript{46}NO^{15}\text{N}_4</td>
<td>oxygen migration</td>
</tr>
<tr>
<td>378.2210 (1.1)</td>
<td>1.05</td>
<td>C\textsubscript{12}H\textsubscript{16}ON\textsubscript{15}\text{N}</td>
<td>C\textsubscript{24}H\textsubscript{20}O^{15}\text{N}_3</td>
<td>pyrrolic (^{15})N loss</td>
</tr>
</tbody>
</table>
Figure 3-4: Fragmentation of 5-NO2 OEP leading to the m/z 522 ion.
is m/z 375. This ion has the elemental composition $C_{24}H_{29}N_3O$, i.e. it is formed by loss of $C_{12}H_{16}N_2O$. The loss of two nitrogens implies that the ion must be formed by cleavage of at least one pyrrolic ring. Such fragmentation processes have previously been observed for substituted porphyrins only as a surface-induced fragmentation process in CIMS [Beato et al., 1989a].

In order to obtain more information on the pathway to the formation of the m/z 375 ion, the MS/MS parent ion spectrum of was obtained (Figure 3-2d). This indicated that the molecular ion was the major parent ion and that the ions m/z 522, 404 and 390 were also parents. The m/z 404 parent supports the idea that an ethyl group from another pyrrolic ring is a part of the fragment that is lost in the formation of m/z 375. The absence of m/z 562, m/z 533, and m/z 550 in the parent spectrum indicates that molecular ions losing OH', NO2, or $C_2H_5^+$ will not fragment directly to give m/z 375 to any significant extent.

The analyses were repeated using $^{13}$C and $^{15}$N isotopically labelled porphyrins. In order to confirm that only one pyrrolic nitrogen is lost to form the m/z 375 ion, $^{15}$N$_4$ labelled 5-NO$_2$ OEP (Figure 3-1) was run by HRMS. A fragment ion appeared at m/z 378.2210 indicating a composition of $C_{24}H_{29}O^{15}N_3$; the loss of $C_{12}H_{16}ON^{15}N$ indicates that the nitrogen of the nitro group and the nitrogen of one of the pyrrole groups is lost. The data for the $^{13}$C$_4$ labelled compound (Figure 3-1) confirm this interpretation. The HRMS spectrum shows an ion at 378.2400 with the composition $C_{21}H_{29}O^{13}C_3N_3$ (Table 3-1). The fragment lost has
the composition $C_{11}H_{16}^{13}CON_2$; the obvious routes to this loss involve the cleavage of one pyrrolic unit, one bridge carbon and the nitrogen of the nitro group.

A proposed scheme for the formation of the m/z 375 ion in the EIMS spectra is shown in Figure 3-5; double arrows indicate rearrangement steps. The HRMS accurate mass assignments indicate that the ions at m/z 522 and m/z 375 occur from losses of $C_3H_5O^+$ and $C_{12}H_{12}N_2O$, respectively (Table 3-1). The m/z 522 ion is the major ion in the parent of the ion at m/z 375. The data for the $^{13}C_4$ labelled 5NO$_2$ OEP indicate that the meso-carbon is not lost in the pathway (Table 3-1). Therefore, one of the carbons lost is likely a $\beta$-carbon from the pyrrolic ring adjacent to the nitro group. If the $\beta$-carbon is lost, then it is most likely that the nitro group will transfer an oxygen to the $\beta$-carbon of the porphyrin, activating the ring towards elimination. It is unlikely that abstraction of a hydrogen atom from the alkyl group by an oxygen of the nitro group is unlikely to be an initial stage in the pathway because m/z 562 does not appear as a parent ion of of m/z 375. Other unlikely possibilities that have not been eliminated by MS/MS studies are the loss of $CH_3^+$ together with HCHO or the loss of $C_2H_5^+$ together with CO. As is shown in Figure 3-5, one possibility is that a pyrrolic ring might cleave forming a ketone, followed by loss of $C_3H_5O^+$ via homolytic cleavage. Rearrangement of the m/z 522 ion can lead to the aromatic ion shown in Figure 3-5. This could eliminate in stepwise fashion with loss of ethyl or methyl forming intermediate ions at m/z 390 and 404 respectively, followed by the expulsion of the remainder of the pyrrolic unit. Alternatively, elimination may also occur in one step with loss of the cyanopyrrole radical to generate m/z 375 ion.
Figure 3-5: Scheme for the ring-scission process of 5-NO_2 OEP.
There are three possible explanations for the differences in the fragmentation pathways between OEP and 5-NO₂ OEP. First, the nitro group may be more labile than the alkyl groups, hence the β-cleavage of the alkyl groups is suppressed because the nitro group fragments more readily than the alkyl groups. This explains the predominance of [M-17]⁺ and [M-46]⁺ ions in the spectrum. This does not, however, explain the α-cleavage of the alkyl moieties or the cleavage of the macrocycle. If the NO₂ does not modify the properties of the macrocycle this does not explain the appearance of the m/z 375 ion. A second possibility is that the NO₂ group causes the macrocycle to become non-planar, thereby significantly changing the fragmentation process.

Structures of nitro-porphyrins obtained with X-ray crystallography clearly show the nonplanarity of the nitro porphyrins [Wu et al., 1991; Shelnutt et al., 1992]. In their characterization of 5-NO₂ OEP using ¹H NMR, Bonnett and Stephenson [1965] showed that the structure is non-planar in solution. The conformation of protonated 5-NO₂ OEP has been shown to be especially distorted owing to both the electrostatic repulsion of the positively charged pyrrolic nitrogens and the steric interactions between the nitro group and the adjacent β-substituents [Meot-Ner and Adler, 1975]. Gong and Dolphin [1984] showed similar results in their UV-visible spectrophotometric characterization of nitroporphyrins; the addition of a nitro group to OEP causes a bathochromic shift and a weakening of all of the visible absorption bands. Theoretical calculations show that a direct correlation exists for the degree of macrocycle distortion and the bathochromic shifts exhibited in the absorption spectrum of porphyrins [Fajer et
at al., 1985]. Using the principles of Goutermans "four orbital" theory [1978] and UV-vis absorption and emission data, Wu et al. studied porphyrins bearing a meso substituent [1991]. They showed that apart from the electronic effects of the substituent itself (the presence of electron-donating groups raises the energy of the highest occupied molecular orbital (HOMO) while electron-withdrawing groups lower the energy of the HOMO), the steric effects were substantial in red-shifting the absorption bands.

We propose that the gas-phase molecular ion of 5-NO₂ OEP retains its distorted shape and that the positive charge may even accentuate this effect. The non-planarity disrupts the conjugated \( \pi \)-bonding system of the macrocycle sufficiently to produce a significant weakening of the bonding system of the macrocycle. This could account for the cleavage of the pyrrolic ring leading to the observance of the m/z 375 ion in the spectrum. It is well known that cyclic tetrapyrroles which are not fully aromatized readily undergo cleavages of pyrrolic units [Boylan, 1969; Budzikiewicz, 1978]. Other evidence is that 5-NO₂-OEP shows anomalous chromatographic behavior in that it is much less polar than OEP in normal phase chromatography (indicating a disruption of the \( \pi \)-system) [Bonnet and Stephenson, 1965].
Analysis of the Divalent Metal Complexes of 5-NO₂ OEP

Spectroscopic studies

There are many reports of systematic studies on the effect of the chelated metal ion and its axial ligands on the spectroscopic properties of porphyrins [Falk, 1964; Gouterman, 1978]. Nappa and Valentine [1978] studied the effect of axial ligands on the visible absorption spectra of zinc porphyrins. Wang and Hoffman [1984] studied the trends in the optical spectra of several biologically important metalloporphyrin enzymes. Spaulding et al. [1975] were the first to show the dependence of certain resonance Raman frequencies on the core size of metalloporphyrins. Parthasarathi et al. [1987] studied trends in the resonance Raman (RR) frequencies of several metalloporphyrin derivatives with an emphasis on the implications for the interpretation of hemoglobin photoproduct RR frequencies.

Stanley et al. [1993] have studied the metal-dependent bands in the IR spectra of a series of divalent metal complexes of OEP as dispersions in cesium iodide pellets. By studying the porphyrins in the solid state, it was possible to use X-ray crystallographic data with more confidence. As a result of this study, it was possible to draw a meaningful correlation between metal-pyrrolic nitrogen bond distances (from X-ray) and precise peak positions of metal dependent bands in IR. The Cₘeso-H bond vibrations (e.g., IR bands ca. 3050 and 1230 cm⁻¹) were the
most conformationally sensitive, as evidenced by the distinctly different IR absorbances for the planar and ruffled (non-planar) forms of OEP.

Anderson et al. [1993] compared the RR frequencies of a series of metal derivatives of OEP and of 5-NO$_2$ OEP. They observed that the metal porphyrin complexes including the Ni(II) undergo changes in conformation in solution. Furthermore it was possible to relate both the planar-nonplanar equilibrium and the position of Raman structure sensitive marker bands to core size. The most sensitive marker band, $v_{10}$, involves the $C_{\alpha}-C_{meso}-C_{\alpha}$ bond vibrations. These relationships stem from the fact that the porphyrin must contract or expand in order to chelate with metals either smaller or larger than the optimum size. The study also indicated that the nitro stretching vibrations are coupled to vibrational modes for the porphyrin macrocycle.

Molecular modeling as well as X-ray crystallographic data were used to determine the core sizes for the complexes [Anderson et al., 1993]. The modelling data also indicated that the repulsions between the nitro group and the adjacent ethyl groups contribute to an increase in the core size. It was reported that the core size is an indicator of planarity; thus the RR frequency shifts can be said to be measure a of the planarity of the macrocycle [Anderson et al., 1993].
Mass spectrometric studies

In the EIMS analysis of octaethylporphyrin (OEP) and its metal complexes, the metal was found to be important in determining the relative abundance of ions in the singly charged region and doubly charged region of the mass spectrum [Beato et al., 1989b]. In contrast, the fragmentation patterns in these regions were not substantially influenced by the nature of the metal ion. The ratio between the summation of doubly-charged fragment ions and the summation of singly charged fragment ions ($\Sigma F^2+/\Sigma F^+$) and a modified Buchler stability index ($S_I$, see chapter 1) for the inserted metal ion was plotted, and the results indicated that there might be some correlation between these parameters [Beato et al., 1989b]. The fragmentation patterns observed were essentially otherwise unaffected by the metal ion; the abundance of singly-charged $\beta$-cleavage fragment ions, $\Sigma F^+$, is relatively unaffected by the metal when compared to the changes in the abundance of doubly charged fragment ions, $\Sigma F^{2+}$.

In the EIMS spectra for the divalent metal complexes of 5-NO$_2$ OEP, the effect of the metal on the ratio $\Sigma F^{2+}/\Sigma F^+$ was not pronounced. The ratio showed no clear correlation when plotted versus $S_I$. The intensities of certain singly charged EI fragment ions appeared to have a significant dependence on the identity of the inserted metal ion. Also, no ions corresponding to the loss of a pyrrole unit are observed in EIMS for any of the metallated 5-NO$_2$ OEPs studied.
In order to more closely study the metal dependence on the singly charged fragment ions of the divalent metallated 5-NO₂ OEPs, MS/MS and HRMS data were obtained. The daughter ion spectra for the metalloporphyrins were examined to determine how the metal alters the fragmentation of 5-NO₂ OEP. The systematic trends in the fragmentations were related to inherent physico-chemical properties of the inserted metal and to spectroscopic data in order to determine if a correlation exists.

The EI MS/MS fragmentation observed for 5-NO₂-OEP was significantly altered by insertion of a divalent metal. The silver complex of 5-NO₂ OEP behaved anomalously and will be discussed separately. The daughter spectra for the molecular ions of the Mg(II), Cu(II), and Co(II) are shown in Figure 3-6 a,b, and c, respectively. Ions at [M-15]⁺ and [M-17]⁺ corresponding to the loss of CH₃⁺ and of OH⁻ were detected for all of the metals, as for metal-free 5-NO₂ OEP. In contrast to the metal-free 5-NO₂ OEP, there was no evidence of pyrrole ring fragmentation for any of the complexes studied. Significant ions at [M-75]⁺ and [M-86]⁺ were noted for all of the metals. Both of the corresponding ions were minor in the daughter spectrum of metal-free 5-NO₂ OEP.

In view of the complexity of the fragmentation pathway for 5-NO₂ OEP, it was necessary to confirm the interpretation of the fragmentation pathways for the metalloporphyrins by HRMS. Thus, Co(II) 5-NO₂ OEP was analyzed by this technique. Unsurprisingly, the [M-15]⁺ and [M-17]⁺ were confirmed as being due to losses of CH₃⁺ and OH⁻, respectively. The [M-75]⁺ ion had a composition of
Figure 3-6: Selected EIMS/MS daughter ion spectra of the M\textsuperscript{**} ions of metallated complexes of 5-NO\textsubscript{2} OEP:  a) Mg(II) complex (m/z 601); b) Cu(II) complex (m/z 640) c) Co(II) complex (m/z 636); d) Ag(II) complex (m/z 684).
Figure 3-7: Formation of the [M-75]⁺ ion for the metallated complexes of OEP.
C_{34}H_{38}N_{4}\text{Co} \text{ (exact mass 561.2381, } \Delta 4.7 \text{ mmu)}, \text{ which confirmed our interpretation of the loss of } C_2H_5^+ \text{ and } NO_2. \text{ The } [M-86]^+ \text{ ion had a composition of } C_{31}H_{33}ON_2\text{Co} \text{ (exact mass 550.2013, } \Delta 0.4 \text{ mmu), which corresponds to a loss of } C_5H_{10}O. \text{ The fragmentations leading to the } [M-86]^+ \text{ ion are interpreted in more detail below.}

The [M-75]^+ \text{ ion resulted from loss of } NO_2 \text{ and } C_2H_5^+ \text{ (Figure 3-7). Note that the molecular ion is not indicated as a radical since this is dependent on the identity of the metal. The daughter spectra of [M-75]^+ \text{ ions had abundant peaks corresponding to } \beta\text{-cleavage. The relative abundance of the collision-induced dissociation (CID) fragment ions of the [M-75]^+ \text{ ions were relatively independent of the inserted metal. This is to be expected because once the nitro group is lost the compound should behave similarly to OEP. This supports the HRMS data for the composition of [M-75]^+. In the parent spectra of [M-75]^+ \text{ ions, the molecular ion was the only detected ion. This suggests that the loss of } NO_2 \text{ and } C_2H_5^+ \text{ are concerted within the time scale of CID, as shown in Figure 3-7.}

The [M-86]^+ \text{ ions resulted from the neutral loss of } C_5H_{10}O. \text{ This provides strong evidence that migration of at least one oxygen is an important step in the fragmentation. Interestingly, parent spectra of the [M-86]^+ \text{ ions showed that molecular ions were the only detectable parent ions. It is not clear how this unusual fragment ion arises. It would seem likely that an oxygen of the nitro group migrates and activates the cleavage of the porphyrin macrocycle. We have tentatively proposed a fragmentation pathway that accounts for the elemental composition of the daughter ion and also results in the generation of a new } 22 \pi\text{-electron macrocycle (Figure 3-8), which could conceivably display mass}
Figure 3-8: Fragmentation of the metallated complexes of 5-NO\textsubscript{2} OEP to give the \([M-86]^+\) ion.
spectrometric behavior similar to OEP. The daughter spectra of the \([M-86]^+\) ions showed intense peaks corresponding to \(\beta\)-cleavage of peripheral ethyl groups. It is possible that the chelated metal ion could be penta coordinate in the \([M-86]^+\) ion. Clearly, the oxygen-activated cleavage of the porphyrin macrocycle is of importance to both metal-free and metal complexes of 5-\(\text{NO}_2\) OEP because the \([M-57]^+\) ion (m/z 522) in the metal-free porphyrin is the product of the loss of a C\(_3\)H\(_5\)O group, which probably is formed via a similar process.

For metal complexes, the loss of 204 to give a ring-cleavage product was not observed. However, the loss of 86 may be the result of a similar pathway to that for m/z 522 (Figure 3-8). The difference is that there is loss of C\(_2\)H\(_5\)^+ (presumably lost via \(\alpha\)-cleavage) in addition to the loss of C\(_3\)H\(_5\)O^-. The reason there is no further fragmentation of the pyrrole ring is probably the templating effect of the chelated metal ion.

One of the more unexpected features of this study is that the fragmentation pathway of silver (II) 5-\(\text{NO}_2\)-OEP is completely different from those of the other complexes studied. The daughter ion spectrum, shown in Figure 3-6d, differs significantly those of the other metals (see Figures 3-6 a,b, and c, for example). The \([M-75]^+\) ion dominates the spectrum with the M\(^+\) ion being the only other ion with a relative abundance greater than five percent. One possibility we have considered to explain this behavior is that silver might assume a monovalent state during the volatilization process; however, we did not detect any ions corresponding to loss of silver which would likely result from such a process. It
should be noted that silver has the largest ionic radius of the metals studied, which could be one factor that causes the anomalous behavior of the ion.

The [M-15]$^+$, [M-17]$^+$, [M-75]$^+$, and [M-86]$^+$ daughter ion relative abundances (\% RIC) have been plotted as a bar graph in order to more clearly determine the metal dependency of each of these ions (Figure 3-9). The [M-17]$^+$ ion abundance seemed to be relatively independent of the metal while the [M-75]$^+$ ion showed a small degree of metal-dependency. The [M-86]$^+$ and the [M-15]$^+$ ions showed the highest degree of dependence on the metal. The relative abundance of the $\beta$-cleavage peak at [M-15]$^+$ was influenced by the presence of various metals, and was inversely related to the intensity of the [M-86]$^+$ peak (Figure 3-9). Evidently, if a metal induces $\beta$-cleavage, the fragmentation pathway leading to [M-86]$^+$ is inhibited.

The relative abundances of these daughter ions were plotted versus the properties of the inserted metal ions to determine whether a significant relationship exists. The properties plotted were the Pauling electronegativity, the ionic radius, the number of $d$-electrons, and the Buchler stability index. The only property that showed a relationship to any of the ion intensities was the number of $d$-electrons, as discussed below. The intensities of the daughter ions were plotted against several of the metal-dependent spectroscopic properties: IR band shifts, UV-visible absorption band shifts (the Soret, $\alpha$, and $\beta$ bands), resonance Raman marker band shifts, and X-ray structural information (core sizes). The IR bands and the Raman lines showed a trend when plotted against the intensity of the [M-86]$^+$ ion, as discussed below.
Figure 3-9: Bar graph showing the metal dependency of the [M-15]⁺, [M-17]⁺, [M-75]⁺, and [M-86]⁺ daughter ions of the divalent metal complexes of 5-NO₂ OEP.
Figure 3-10a shows that an apparent trend in the intensity of the [M-86]+ ion exists for the metal complexes studied. The plot indicates that the intensity of the [M-86]+ ion decreases as the number of d electrons for the inserted metal increases. Note that Mg was placed with Zn since the d⁰ and d¹⁰ configuration should act similarly when the d orbital effect on the ring bonding system is considered. The plot in Figure 3-10b shows that a similar trend exists for the metal-sensitive IR band ca. 3050 cm⁻¹ for OEP [Stanley et al., 1993]. This band corresponds to the bridge carbon-hydrogen stretch. The 3050 cm⁻¹ band was chosen since it is the most sensitive of the five metal-dependent IR bands of OEP and it is one of the bands that is conformation-dependent. However, the general trend is nearly the same for the other bands at 1230 and 920 cm⁻¹ [Stanley et al., 1993]. Further studies are needed for additional compounds to verify the trends and to determine the role of the metal in altering the fragmentation pathways.

Figure 3-11a shows that a trend exists between the metal dependent IR band at 3050 cm⁻¹ for OEP and the % RIC for the [M-86]+ ion. Figure 3-11b indicates the trend is similar for the ν₁₀ structure-sensitive Raman marker band. These relationships indicate that a broad trend exists between the bond strengths of the Cₐ⁻Cₐ⁻Cₐ and Cₐ⁻H bonds, which may reflect on the conformation of the macrocycles. More data are needed in order to fully understand the implications of the observation.
Figure 3-10: Plots of: a) the [M-86]$^+$ daughter ion intensities vs. the number of $d$-electrons for divalent metals; b) the IR metal dependent band at 3050 cm$^{-1}$ vs. the number of $d$-electrons for divalent metals.
Figure 3-11: Plots of: a) the metal-dependent IR band at 3050 cm\(^{-1}\) vs. the metal-dependent daughter ion at [M-86]\(^+\) b) the metal-dependent Raman marker band, \(v_{10}\) vs. the metal-dependent daughter ion at [M-86]\(^+\).
Conclusions

The fragmentation of 5-NO₂ OEP was found to be significantly different from that of OEP. Loss of OH⁻ and α-cleavage become important modes of fragmentation when a nitro substituent is attached to the bridge carbon position of OEP as opposed to the normally encountered β-cleavages of OEP. MS/MS and HRMS data provides evidence for pyrrole ring scission in 5-NO₂ OEP. In the proposed mechanism, one of the oxygens of the nitro group attacks the β-carbon of the adjacent pyrrole rings leading to the loss of C₃H₅O⁻ from the molecular ion (m/z 579) to form an ion at m/z 522, as was confirmed by the HRMS analysis. The parent ion scan of the pyrrole scission fragment ion at m/z 375 showed the m/z 522 ion as the major fragment ion leading to ring scission. The m/z 522 ion undergoes a series of rearrangements resulting in the loss of a pyrrole unit. Previously, ring cleavage of non-substituted porphyrins has been observed only as a surface-induced phenomenon in CIMS. We have attributed the ring-cleavage process to the non-planarity exhibited by the 5-NO₂ OEP. In the nonplanar conformation the aromatization of the macrocycle is weakened and the porphyrin fragments to lose a pyrrole unit. This is consistent with previous findings [Budzikiewicz, 1978] in which tetrapyrroles with non-aromatic meso-carbons readily form ring-scission fragments.

With the insertion of a divalent metal (Mg(II), Ni(II), Cu(II), Co(II), Zn(II), Ag(II), or Pd(II)) into 5-NO₂ OEP, the modes of fragmentation become significantly altered. A peak at either [M-75]⁺ or [M-86]⁺ was the base peak in the daughter
spectra of the $M^+$ ions of all of the metallated species. The $[M-75]^+$ was attributed to loss of NO$_2$ and C$_2$H$_5^+$. The $[M-86]^+$ peak involved oxygen migration in a process similar to that described for the non-metallated 5-NO$_2$ OEP. The fragmentation involves the loss of C$_3$H$_5$O$^+$ analogously to 5-NO$_2$ OEP; but the loss of an additional C$_2$H$_5^+$ also occurs. The ring-cleavage phenomenon was not observed for any of the metal complexes studied. We have attributed this effect to the stabilization incurred by templating of the metal ion with the macrocycle.

The intensity of the ions in the daughter spectra of the $M^+$ ions showed a marked dependence on the identity of the chelated metal ion. The $[M-86]^+$ and the $[M-15]^+$ ($\beta$-cleavage) ions were observed to be the most metal dependent of the fragment ions. Other significant ions at $[M-75]^+$ and $[M-17]^+$ (loss of OH$^+$) showed a much lesser degree of metal-dependence. The silver II complex behaved anomalously as the $[M-75]^+$ ion was the predominant ion appearing in the daughter spectrum with all other fragment ions having an intensity less than 5 percent.

The intensity of the $[M-86]^+$ ion was plotted against the number of d-electrons for the inserted metal atom. The $[M-86]^+$ intensity was found to increase as the number of d electrons of the inserted metal decreased. A similar trend was shown for the metal-sensitive IR band at 3050 cm$^{-1}$ for OEP. Plots of the 3050 cm$^{-1}$ IR band and of the the $\nu_{10}$ Raman band versus the $[M-86]^+$ daughter ion showed a similar trend. The 3050 cm$^{-1}$ band corresponds to the C$_\text{meso}$-H stretch, while the $\nu_{10}$ Raman band corresponds to the $C_\alpha$-C$_\text{meso}$-$C_\alpha$. 
vibration. Both of these bands have been shown to be sensitive to the planar-nonplanar conformations of metallated porphyrins.
CHAPTER 4
TANDEM MASS SPECTROMETRIC STUDIES OF PHOTOSENSITIZATION REACTIONS OF PORPHYRINS

Introduction

Thermospray ionization/tandem mass spectrometry TSP/MS/MS has been employed to monitor photochemical reactions occurring at biologically relevant wavelengths. A high intensity xenon arc lamp (300 W) affords a high yield of photoproducts, while MS/MS data provide information to elucidate possible structures of these products. The potential use of such instrumentation is in the investigation of porphyrin photosensitization reactions with classes of biomolecules such as amino acids, nucleic acids, and short chain peptides. The present study reports the on-line monitoring of the photosentization of tryptophan (Trp) by hematoporphyrin (HP). Off-line experiments were performed for mixtures of Trp with HP, tetramethylpyridilporphyrin (TMePyP), and tetrasulphonatophenylporphyrin (TSO₃PP). The photochemistry cell has been constructed in our laboratory and is evaluated by comparison to off-line experiments. Pertinent background areas for this research include photochemistry, photosensitizing reactions, porphyrin photosensitizers,
photodynamic therapy, and aqueous-phase chemistry on-line with mass spectrometry as detailed below.

Photochemistry

Photochemical processes are of utmost importance to all life on earth. Photosynthesis is the most significant example of tapping the sun's energy; virtually all of our available energy resources can be ascribed to conversion of sunlight into chemical potential energy. Furthermore, this process allows the conversion of abiological carbon into organic compounds. Photochemistry has played a crucial role in organic synthesis [Turro, 1978]. The medical applications of photochemistry have become increasingly valuable in recent years [Neckers, 1989]. Photodynamic cancer therapy, discussed below, is a prime example.

Photosensitization Reactions

A schematic representation of a photosensitization reaction is shown in Figure 4-1. Photosensitization reactions occur when a sensitizing molecule absorbs light, achieving a reactive state which reacts with substrate molecules. The excited state of most molecules is the singlet state. In order for the molecule to be an effective sensitizer, intersystem crossover (change in the spin of the
Figure 4-1: Schematic representation of the type I and type II photosensitization reactions (Sens= Sensitizer).
excited electron) to the triplet state must occur, since the triplet state is longer-lived and more reactive than the singlet state [Foote, 1984].

There are two types of photosensitizing reactions [Foote, 1976]. In the type I reaction, the sensitizer interacts directly with the substrate, inducing radical formation either by hydrogen abstraction or by electron transfer. The substrate radical is then oxidized upon reaction with oxygen. In the type II reaction, the sensitizer transfers its excitation energy to ground state (triplet) oxygen. The oxygen is converted to the more reactive excited singlet state, which subsequently reacts with the substrate molecule.

One of the characteristics of the type I reaction is that the reaction rate is dependent on the concentration of oxygen [Foote, 1976]. This is because oxygen and the substrate are always in competition for the excited triplet state sensitizer. This process is shown schematically in Figure 4-1.

Photosensitization reactions of interest in many areas of photochemistry and photobiology have been extensively researched. Other areas of interest have also been widely studied. Extensive reviews on the damages induced in nucleic acids by photosensitization reactions have appeared [Wang, 1976; Piette and Decuyper, 1986]. Neckers [1989] has reported on the mechanisms of phototherapy for neonatal jaundice. Recently, Garcia [1994] reviewed the implications of photosensitizing reactions in the degradation of aquatic phenolic pollutants. The area covered in this dissertation is that of the photodynamic therapy (PDT) for cancer.
Porphyrin Photosensitizers

Porphyrins, in general, are excellent photosensitizers due to their high molar absorptivities (e.g. $\epsilon > 5 \times 10^5$ L mol$^{-1}$ cm$^{-1}$) and their high quantum yields for intersystem crossover to the triplet excited state. For instance, hematoporphyrin and its derivatives have quantum yields near 0.9 for the singlet to triplet transition [Cannistraro et al., 1983].

It has long been known that certain porphyrins cause photosensitivity in man; the photosensitivity associated with malfunctions of porphyrin metabolism (e.g., the porphyrias) is attributed to the presence of photoactive porphyrins in the skin [Blum, 1941]. Dougherty et al. [1978], working at the Roswell Park Cancer Institute, first reported the use of the hematoporphyrin derivative (HPD) in the treatment of patients with recurrent skin cancer. Thousands of cancer patients have since been successfully treated with Photofrin $^\text{TM}$, the commercial product that is otherwise known as HPD [Kessel et al., 1987; Grossweiner, 1994]. Photofrin has been approved for phase II and phase III clinical trials in the United States [Grossweiner, 1994].

Photofrin is a porphyrin porfimer made by reacting hematoporphyrin with sulfuric and glacial acetic acid and then washing and neutralizing the resultant brown powder with NaOH by the method first reported by Lipson in 1961 (Figure 4-2). This hematoporphyrin derivative (HPD) is a complex polymeric mixture containing tumor-localizing components (DHE) and non-localizing components,
Figure 4.2: Synthesis of hematoporphyrin derivative (HPD) from hematoporphyrin
as well as hematoporphyrin (HP) itself and its dehydration products. Musselman et al. [1988] have analyzed HPD by fast atom bombardment mass spectrometry and have shown that polymers up to the 8-mer exist. Both HP and DHE absorb weakly at their highest-wavelength band of ca. 630 nm; this is significant since the wavelength of irradiation used in photodynamic therapy (PDT) is typically above 600 nm to maximize transmittance through tissue. It is clear that the development of new sensitizers with well-defined compositions and higher absorbances at longer wavelengths would be advantageous.

Synthetic water soluble anionic and cationic porphyrins have recently received considerable attention as potential PDT agents since they have absorption bands well above 600 nm [Oseroff et al., 1990; Ruck et al., 1992; Ando et al., 1993]. The chlorins [Nelson et al., 1990] and the phthalocyanines [Van Lier, 1990; Olenick et al., 1993] have also recently been suggested as potential PDT agents due to their long-wavelength absorption bands.

Clinical Photodynamic Therapy (PDT)

Photodynamic cancer therapy involves the injection of a suitable sensitizer into the bloodstream of the cancer patient and illumination of the affected area with light, typically from a laser [Dougherty et al., 1978; Parker, 1990]. The sensitizer is a dye that must be able to be localized in the area of a tumor, and must have a high triplet state quantum yield. Some photosensitizers can act
directly through the excited triplet state (type I), but most (including Photofrin) act indirectly with a spin-conserving intermolecular energy transfer to triplet oxygen (type II) [Sonoda et al., 1987]. Singlet oxygen is produced and is believed to react with biomolecules such as amino acids in proteins and lipids that compose cell structures (Figure 4-3). Neither the exact mechanism of tumor cell destruction nor the critical sites of subcellular damage are known precisely. Some evidence suggests that damage to the mitochondria with subsequent inhibition of cell respiration is the cause of death [Henrickson, 1992]. The localization and transport of porphyrins that occur in vivo are affected by the physio-chemical properties of the porphyrin. Hydrophobic porphyrins have a high affinity for albumin and other proteins and have high partition coefficients with those proteins [Korbelik, 1993]. The binding of the porphyrin photosensitizer to proteins seems to play an important role in the accumulation of the sensitizer in the tumor cells [Milanski et al., 1987]. The clinical aspects of photosensitizer delivery are reviewed in detail in the proceedings of the Clayton Foundation [Gomer, 1987a,b].

On-line Solution Chemistry

The interfacing of a "chemical reactor" with mass spectrometry allows the monitoring of solution chemistry. The inherent speed and selectivity of on-line MS/MS techniques offer the ability to directly monitor reactants, short-lived intermediates, and products of reactions. The thermospray interface permits on-
Figure 4-3: Porphyrin type II photosensitizing reactions of biomolecules
line monitoring of such a reactor at relatively high flow rates (1-2 mL/min) [Vestal, 1983, 1984]. Thermospray ionization provides simple mass spectra, typically dominated by [M+H]^+ or [M-H]^- ions, even for polar, involatile, and thermally labile compounds. Tandem mass spectrometric techniques allow for structure elucidation of the ions produced in the thermospray process [Volk et al., 1992].

Previous work in the research group developed a technique for the on-line monitoring of an electrochemical cell and an enzymatic reactor by TSP/MS/MS [Volk et al., 1992]. Unique insights into reaction pathways were achieved with the coupling of these reaction cells with mass spectrometric detection. The ability to monitor intermediates and reactants as a function of such parameters as electrochemical potential and enzymatic reaction time constitutes a powerful approach to exploring the thermodynamics and kinetics of chemical reactions.

The first applications of photolysis-thermospray-mass spectrometry have recently been demonstrated. At the 1992 meeting of the American Society for Mass Spectrometry, Ballard and Grinberg showed the on-line detection of the photolytic cleavage of biologically active compounds by mass spectrometry. Lurie et al. [1992] recently showed the use of on-line photolytic derivatization of forensic drugs with mass spectrometric detection, in which photolysis increased the extent of ionization of certain compounds which are otherwise difficult to analyze.
Thermospray Ionization

The thermospray interface is well-suited to the monitoring of solution-phase reactions because of the relatively high flow rates to be used (and the small dead times, typically < 1 s, which result), as well as the capability for ionization of polar and involatile molecules [Vestal, 1983, 1984]. Arpino [1992] has written an exhaustive review on the LC/MS applications of thermospray.

Instrumentation

The interface consists of a vaporizer probe and thermospray ion source, equipped with a thoriated iridium filament, discharge electrode, liquid nitrogen solvent trap, and pump-out line to a 300 L/min mechanical pump. The thermospray ionization source is shown in Figure 4-4a.

The TSQ 45 tandem mass spectrometer allows for fragmentation of the [M+H]+ and [M-H]- ions typically produced to give structural information on intermediates and products of the photochemical reactions. The detection limits for TSP/MS/MS are typically in the low ng levels for compounds amenable to thermospray.
Figure 4-4: The a) thermospray interface and b) schematic of the thermospray ionization process.
(a) 

ION SAMPLING CONE 
IONS 
ELECTRON BEAM 
VAPORIZATION PROBE 
HEATER 
VAPORIZATION COUPLING 
MOUNTING POST 
BLOCK TEMPERATURE 
TO TRAP B MECHANICAL PUMP 

(b) 

TO MASS SPECTROMETER 
VAPORIZATION 

PUMP 
H₂O 
HEAT 
HEAT 
LIQUID 
H₂O 
H₂O 
NH₄⁺ 
NH₄⁺ 
H₂O
Theory

The thermospray ionization process occurs during the rapid vaporization of a volatile buffer solution without any means of external ionization (Figure 4-4b) [Vestal, 1983, 1984]. Usually the buffer is ammonium acetate; as the droplets are formed, there is a statistical distribution for the accumulation of net positive or net negative charge in each droplet (excess NH₄⁺ or Ac⁻ ions, respectively). The droplets quickly vaporize and develop a high enough charge to permit the vaporization of larger ions or small ion clusters. These ions are then drawn into the mass spectrometer through the ion sampling cone. This ionization method yields [M+H]⁺ and [M-H]- (or [M+NH₄⁺]⁺ and [M+Ac⁻]⁻ ions for most nonvolatile, polar compounds not amenable to conventional chemical ionization (CI) mass spectral analysis.

External means of ionization with the filament or discharge electrode provide for ionization of samples in nonaqueous solutions or those lacking buffer. In the "filament-on mode" conventional positive and negative CI analyses can be carried out using the vaporized mobile phase as the reagent gas [Vestal, 1984].

The thermospray desolvation process produces backpressures that can exceed 1000 psi. When the solvent quickly evaporates within the 0.015 mm ID capillary, the supersonic gas expansion produces the excessive pressure. The backpressure of the thermospray ionization process proved problematic in the design choice of a photochemical cell, as discussed below.
On-line Flow Cell Development

The design of a photolysis cell suitable for coupling to a tandem mass spectrometer via a thermospray interface is described. The construction of a photochemical cell was necessary because there are no commercially-available photolysis cells which can both withstand the high backpressures generated in the thermospray ionization process and provide sufficient volume to give residence times on the order of one minute.

Commercially Available Flow Cells

In the previous report of on-line photolysis with thermospray ionization mass spectrometry [Ballard and Grinberg, 1992], a commercially-available knitted, open-tubular reactor (KOTR) was used. The cell is constructed from PTFE and is used in LC applications requiring a minimum of band broadening effects. The reaction cell comes equipped with a 254 nm mercury vapor lamp (8 W) and is marketed as PHRED™ (photochemical reactor for enhanced detection, Aura Instrumentation, Staten Island, NY). The cell is rated to only 350 psi, but Ballard and Grinberg [1992] reported that approximately 50 injections were possible before the cell would burst.

Since the goal of this research was to examine photosensitizing reactions at visible wavelengths using a high-power Xe arc lamp (300 W) as the light
source, the use of KOTRs was precluded. The high heat output of the lamp would have probably caused a considerable decrease in the lifetime of the KOTR cell.

Since the minimization of band broadening effects was also a concern, the use of fused silica tubing with UV-visible-transparent coatings (CElect™-UVT, Supelec, Inc., Bellefonte, PA) was also examined. The small volume of the capillary tubing and the high thermospray flow rates result in a residence time within the cell of too short a duration, thus the flow cell utilized in this study has been constructed in our laboratory (see below).

Residence Time Requirements

The conversion efficiency is a function of the photon flux through the photolysis cell, the residence time of the analyte in the cell, the molar absorptivity, and the quantum yield of the photochemical process. The residence time depends primarily on the volume of the cell since thermospray flow rates typically vary only between 0.5 and 2 mL/min. Ballard and Grinberg [1992] reported residence times of between 2.9 and 5.4 min in the KOTR, which was sufficient to convert benzophenone to detectable quantities of benzopinocal in their study. With the high intensity of the Xe arc and the high absorption coefficients of the porphyrins employed in our study, the residence time was chosen to be 1 min.
The response time is dependent on the cell volume as well as the dead volume between the reactor and the ionizer and the flow rate. Past studies performed with an electrochemical reactor placed directly before the thermospray probe showed a minimum response time of 500 ms [Volk et al., 1992]. In the present study, it was not possible to place the photochemical reaction cell as close to the thermospray probe; as a result, longer response time were anticipated.

**High Pressure Photochemical Reaction Cell**

A simple cylindrical glass reactor was constructed from Pyrex 7740 1/4" OD tubing (ID: 4 mm). Glass was used since it absorbs strongly in the UV; the direct absorption of light energy by the tryptophan (which absorbs only in the UV) is thus minimized. Several different types of ferrules were tried to determine which ones could best permit the assembly to withstand high backpressures without leaking. Graphitized Vespel™ ferrules proved to give the best seal of the ferrules tested. Sanding of the ends of the glass near the ferrule seal was performed, and increased the maximum pressure at which the fittings would slip off (about 800 psi, the average backpressure encountered in thermospray). However, the Swagelok fittings would also slip off during fluctuations in the thermospray backpressure. In order for the cell to more reliably withstand the high backpressures of the thermospray interface without having the Swagelok
fittings slip off, a brace was constructed to hold them in place as is shown in Figure 4-5. The braced cell has been able to withstand backpressures as high as 2000 psi.

Experimental

All data presented here were collected using a Finnigan TSQ 45 mass spectrometer with a Vestec thermospray ion source installed. The buffer solution consisted of 0.1 M ammonium acetate solution (pH 7), or 0.1 M ammonium acetate with the pH adjusted to 10.4 by the addition of concentrated ammonium hydroxide. The 300 W Xe arc lamp was operated at 17 A for all of the photochemical experiments.

The pumping was provided by an HPLC syringe pump (ISCO, LC-2600) that has flow rates ranging from 1 μL/min to 7 mL/min and a 500 mL capacity. The flow rate was set at 1 mL/min for all experiments reported here. The pump provides pulse-free delivery up to an upper pressure limit of 3,700 psi. A loop injector was used to inject samples into the mobile phase flow. A 2 μm filter element was placed after the injector to minimize problems with clogging of the thermospray probe.

All mass spectra were obtained in the filament-off mode with a source temperature of 250-260 °C and a probe tip temperature of 205 °C. The collision energy for all of the MS/MS experiments was 30 eV with an argon pressure of 1.4 mTorr in Q2. Tuning for the MS/MS experiments was performed by injecting 1 ml
Figure 4-5: Diagram of the high-pressure flow cell.
of a tryptophan solution. The electron multiplier was set at -1200 V for the MS experiments and -1600 V for the MS/MS experiments.

On-line Setup

A schematic of the on-line photolysis system is shown in Figure 4-6. The setup consists of an Isco LC-2600 syringe pump, a loop injector, a photochemical reactor, a Vestec thermospray probe and ion source, a Finnigan TSQ 45 mass spectrometer, a liquid nitrogen solvent trap, and a 300 L/min mechanical pump.

The photochemical reactor was designed in our laboratory and consists of a fan-cooled 300 W xenon arc lamp and the flow cell described above. The flow cell is held within a fan-cooled aluminum housing, which was designed to dissipate heat and to reflect light towards the region of the flow cell. The arc lamp and the aluminum housing are contained in a plexiglass box that has been painted black in order to keep stray light from entering or exiting the photochemical reactor. A window was included in the top of the housing so that the alignment of the photochemical reaction cell with the light beam could be observed. The box is vented to a hood to remove ozone and excessive heat that are generated by the lamp.
Figure 4-6: Schematic diagram of the on-line photochemistry TSP/MS/MS system.
Off-line Setup

The setup for the off-line photochemistry experiments is shown in Figure 4-7. The system is composed of a Xe arc lamp, an water-filled infrared (IR) filter, a focusing lens, a magnetic stirrer, and a 100 mL three-neck flask that is fitted with a thermometer and a cold finger. The IR filter removes most of the heat from the light-beam. The lens is focussed toward the center of the flask to maximize the photon flux through the solution. A stir bar provides solution mixing while the cold finger provides temperature stabilization (tap water is passed through the cold finger, maintaining the solution at a temperature of 27° C). The solution is sampled through the unused 3rd neck of the flask as the solution is photolyzed.

Compounds

The compounds employed in this study were tryptophan (Trp), hematoporphyrin (HP), meso-tertramethylpyridylporphyrin (TMePyP), and meso-tetrasulphonatophenylporphyrin (TSO₃PP). The structures are shown in Figure 4-8.

Hematoporphyrin was chosen since several aspects of its photosensitizing reactions have been widely researched [Dubbelman et al., 1982; Smith, 1985; Truscott, 1986; Smith, 1992]. Since one of the initial goals of this research was to detect intermediates of the photosensitization process, it was especially
Figure 4-7: Schematic diagram of the off-line photochemistry setup.
Figure 4-8: Compounds studied.
interesting that Smith [1992] reported the presence of long-lived transients of HP in aqueous solutions that had been flash-photolyzed. Hematoporphyrin is completely soluble in aqueous solutions only at elevated pHS (> 9), with aggregation occurring at lower pHS [Smith, 1985]. Tryptophan was chosen since it is one of the most reactive of the amino acids toward photosensitizers and is often used as a singlet-oxygen trap [Benassi et al., 1967]. The pathways of tryptophan photosensitization reactions in various solvent systems have been previously examined [Savige, 1971]. Also, the photosensitization of Trp by HP has already been studied by Sconfienza et al. [1980]. In our initial thermospray analysis of various substrate molecules, Trp proved to be highly amenable to thermospray ionization with its ion signal consisting almost entirely of the [M+H]⁺ ion (m/z 205). It was also of initial interest to examine the photosensitizing reactions of guanine since its photochemistry has been widely studied. Guanine is the most reactive of the nucleic acids and photosensitization of DNA results in cleavages mostly at the guanine residues [Pasternack et al., 1983]. Guanine is known to interact directly with porphyrin photosensitizers in aqueous solution by extensive overlap of the π-systems of the molecules in a stacking arrangement [Pasternack et al., 1985; Le Nouen et al., 1989]. Unfortunately, in the initial characterization, guanine proved difficult to analyze by thermospray and caused severe problems with clogging of either the thermospray probe or the in-line filter.

The synthetic porphyrins TMePyP and TSO₃PP were chosen since both are highly water soluble over a large pH range. Also, both have been extensively
investigated recently as potential PDT agents [Harriman et al., 1983; Pasternack et al., 1985; Le Nouen et al., 1989; Ruck et al., 1992; Ando et al., 1993].

Methods

For the on-line experiments, the buffer solution consisted of 0.1 M ammonium acetate solution with the pH adjusted to 10.4 by the addition of ammonium hydroxide. $5.0 \times 10^{-4}$ M solutions of tryptophan (Trp) and hematoporphyrin (HP) were prepared in the buffer. Immediately prior to injection, 1.5 mL of Trp solution was spiked with 0.5 mL of the HP, giving concentrations of $3.8 \times 10^{-4}$ M Trp and $1.3 \times 10^{-4}$ M HP. Solutions containing only Trp and HP were diluted to the same respective concentrations prior to injection. One milliliter of each solution was injected manually with a loop injector. The flow rate for all experiments was 1.0 mL/min, which corresponds to an ideal (no band-broadening) residence time of 1.0 min for the cell. Injections were made with the light on and with the light off for all three of the solutions (Trp, HP, and Trp/HP).

For the off-line experiments, the same buffer as above was used in all but the final trial. A $5.0 \times 10^{-4}$ M Trp solution was prepared both in the pH 10.4 buffer and in the pH 7 ammonium acetate solution. Porphyrin solutions were prepared at the following concentrations: HP, $3.0 \times 10^{-4}$ M; TSO$_3$PP, $4.3 \times 10^{-4}$ M; TMePyP, $1.3 \times 10^{-4}$ M; TMePyP (pH 7), $1.3 \times 10^{-4}$ M. Seventy-five milliliters of the Trp solution were put into the three-neck flask which was then spiked with 5 mL of the
individual porphyrin solutions. Aliquots were taken from the solution with a 250 µL disposable pipet, starting with an aliquot immediately after mixing of the solution but before the lamp was turned on. Subsequent aliquots were drawn at intervals during the photolysis of the solution up to a time of one hour. The solutions were injected (100 µL) into the TSP/MS within one hour after the end of the photolysis experiment.

**Results and Discussion**

**Initial TSP Analysis of Porphyrins**

In our initial analysis by TSP/MS of the porphyrins employed for this study, it was discovered that the compounds were not detectable. External ionization techniques (filament on, discharge) coupled with the addition of organic modifiers to the solutions were attempted. Mass spectra of both the positive and negative ions were examined with as much as 50 µg of porphyrin injected. No signal corresponding to M⁺, [M+H]⁺, M⁻, or [M-H]⁻ ions were detected for any of the porphyrin compounds studied. Barely detectable ions below m/z 200 were observed in some cases, but did not correspond to apparent (expected) porphyrin fragments. The low mass ions occurring in the porphyrin TSP mass spectra were attributed to impurities in the samples. Even though the porphyrins were undetectable by TSP, it was determined to continue the study by detection of the tryptophan substrate and its photosensitization reaction products.
The flow characteristics of the flow cell were examined before the on-line photochemistry experiments were attempted. The flow-cell exhibited problems of severe band broadening effects (i.e., an injection led to a ten minute duration of tryptophan [M+H]^+ ion (m/z 205)). One problem resulting from the band broadening was that the analyte solution became seriously diluted. Other problems associated with the delay time between the photochemical event and subsequent thermospray ionization and mass spectrometric detection are discussed in the next section (in relation to off-line photochemistry).

Figure 4-9a shows the normal thermospray mass spectrum of Trp (MW 204). The spectrum exhibits almost exclusively the [M+H]^+ ion at m/z 205. Low intensity ions at m/z 222 (the [M+NH4]^+ adduct) and m/z 161 ([M+H-CO2]^+) were detected. When the same solution of Trp was injected with the lamp on (Figure 4-9b), additional low intensity ions were observed at m/z 145, 177, 201, and 218. In addition, the ion at m/z 161 was significantly increased in intensity.

Figure 4-10a shows the spectrum for the Trp/HP mixture with the lamp off. The ions at m/z 145, 161, 201, and 218 were detected but were much lower in intensity than in the absence of HP but with the light on Figure 4-9b. The peak at m/z 100 was attributed to a low mass impurity in the HP. In the TSP mass spectrum of HP (not shown), m/z 100 was the base peak, although it was barely detectable even at large concentrations.
Figure 4-9: Thermospray mass spectrum of tryptophan with the on-line setup: a) light off; b) light on.
Figure 4-10: Thermospray mass spectrum of the hematoporphyrin/tryptophan mixture with the on-line setup: a) light off; b) light on.
Relative Abundance

% Relative Abundance

M+H+ [M+H]+

145
218
235
265
50.0-
The spectrum for the Trp/HP mixture is shown in Figure 4-10b for the lamp on experiment. The ions at m/z 145, 201, and 218 were dramatically increased in intensity, with numerous other ions also detected. The peak at m/z 205 for the Trp [M+H]^+ ion accounted for only 18% of the reconstructed ion current (RIC). A reduction in the absolute intensity of the m/z 205 ion to 19% of its value in the light-off spectrum was observed, indicating almost 80% conversion of Trp to other species.

The daughter spectrum of the [M+H]^+ ion, m/z 205, of Trp (Figure 4-11) shows the expected daughter ions at m/z 188, 170, 159, and 118 corresponding to neutral losses of NH₃, NH₃+H₂O, HCOOH, and C₃H₅O₂N, respectively.

Daughter spectra were obtained for the main photoproduct ions at m/z 145, 201, and 218 that occurred for the illuminated Trp/HP mixture (see below). These spectra suggest quinazoline derivatives as the main photoproducts. This is different from the generally known biological oxidation pathway to the kynurenine derivatives; however, Savige [1971] has previously proposed that tryptophan is photooxidized to the quinazolines when significant concentrations of ammonia are present in the solution.

In Figure 4-12, the daughter ion spectrum of the m/z 218 ion is shown. The proposed structure is 4-(2-amino-2-carboxyethyl) quinazoline. The dominant ion in the daughter spectrum appears at m/z 145 and can be accounted for as the protonated 4-methyl quinazoline ion (see structure, Figure 4-14). The ions at
Figure 4-11. Daughter ion spectrum of the [M+H]+ ion of tryptophan, m/z 205.
Figure 4-12: Daughter ion spectrum of the hematoporphyrin/tryptophan mixture photoproduct at m/z 218 with the proposed structure of 4-(2-amino-2-carboxyethyl) quinazoline shown.
Relative Abundance

\[ \text{[M+H]}^+ \]

\[ 218 \]

\[ 145 \]

\[ 128 \]

\[ 118 \]

\[ 88 \]

% Relative Abundance
m/z 155 and 183 can be attributed to neutral losses of NH₃+HCOOH and NH₃+H₂O, respectively.

In Figure 4-13, the daughter spectrum of the m/z 201 ion is shown. The proposed structure is 4-(2-carboxy-1-ethenyl) quinazoline. Ions at m/z 183 (loss of H₂O) and 155 (loss of HCOOH) are present. However, an ion corresponding to loss of H₂O+NH₃ was not detected. This corresponds to the proposed structure in which the amine group is no longer present.

The daughters of m/z 145 are shown in Figure 4-14, with 4-methyl quinazoline as the tentatively proposed structure. The m/z 128 ion could be accounted for by loss of NH₃ through a ring opening process. The derivatives of quinazoline were not available to verify our proposed structures.

Off-line

After the completion of the on-line experiments, the off-line setup was designed in order to ascertain if the on-line system afforded the detection of short-lived intermediates of the photochemical reactions. The important problem from the standpoint of detection of short-lived intermediates is that the photoproducts have a time delay from illumination within the cell until ionization and detection of the photosensitization products.

A comparison between the on-line and the off-line thermospray mass spectra of the Trp/HP mixture (Figure 4-15) reveals that the two photochemical
Figure 4-13: Daughter ion spectrum of the hematoporphyrin/tryptophan mixture photoprodct at m/z 201 with the proposed structure of 4-(2-carboxy-1-ethenyl) quinazoline shown.
Figure 4-14: Daughter ion spectrum of the hematoporphyrin/tryptophan mixture photoprotein at m/z 145 with the tenatively proposed structure of 4-methyl quinazoline shown.
Figure 4-15: Thermospray mass spectra of the hematoporphyrin/tryptophan mixture photolysis products: a) off-line; b) on-line.
setups lead to the same photoproducts. The intensities of the ions vary; however, the major ions in both spectra are the same (m/z 145, m/z 161, m/z 175, m/z 201, m/z 205, m/z 218, and m/z 236). It was found that if the solutions from the off-line photolysis were allowed to sit overnight, the spectrum displayed many more ions due to degradation of the photoproducts. However, when the off-line solutions are run within two hours, the results are comparable to the results from the on-line setup.

In order to determine the time-dependence of the photosensitization reactions off-line, the Trp/HP, Trp/TSO₃PP (at pH 10.4), and Trp/TMePyP (at both pH 7 and pH 10.4) mixtures were run with aliquots being taken at intervals during the reaction up to one hour. The plot in Figure 4-16 shows the integrated intensity of the [M+H]+ ion (m/z 205) of tryptophan vs. time for the three porphyrins studied. At least one-third of the tryptophan is photosensitized within the first minute for all three mixtures. After one hour > 99% of the tryptophan has been photo-oxidized in all three cases.

The photoproduct mass spectra are very similar for the three porphyrins studied, as indicated in the spectra shown in Figure 4-17. The spectra shown are for the aliquots taken after 24 min of photolysis. The three main product ions identified as quinazolines from the on-line MS/MS experiments (m/z 145, m/z 201, and m/z 218) are dominant ions in all four of the photoproduct spectra. For TMePyP/Trp (equimolar) mixtures, the products were the same at both pH 7 and
Figure 4-16: Plot of m/z 205 peak intensity vs. time for the porphyrin/tryptophan mixtures
Figure 4-17: Thermospray mass spectra for the photolysis products of the porphyrin/tryptophan mixtures: a) HP/Trp at pH 10.4; b) TSO₃PP/ at pH 10.4; c) TMePyP/Trp at pH 10.4; d) TMePyP/Trp at pH 7.
pH 10.4; the biologically known oxidation pathway to kynurenine and its derivatives was not induced by lowering the pH to 7.

**Conclusions**

Visible-wavelength photochemistry on-line with TSP/MS/MS for the detection and structure elucidation of tryptophan (Trp) photosensitization reaction products has been demonstrated. A 300 W arc lamp provided sufficient light intensity to convert approximately 80% of Trp to photoproducts in the presence of hematoporphyrin (HP) when the reaction was run on-line. Hematoporphyrin was found to increase 50-fold the abundance of photoproducts observed. Daughter ion spectra of the main photoproduct ions at m/z 145, m/z 201, and m/z 218 led to the proposal of quinazoline derivatives as the structural class of the photoproducts. This is in accordance with previous observations that Trp reacts to give quinazolines when photosensitized in the presence of ammonia.

The self-constructed on-line flow cell was able to withstand the high thermospray backpressures (800-1000 psi). However, the cell exhibited severe band broadening effects. A comparison between the on-line and off-line photochemistry setups was performed in order to determine whether the on-line setup afforded the detection of short-lived intermediates. The results show that the on-line flow cell suffers from too much band broadening and must be reduced
in volume and placed much closer to the ionization source in order to have a chance at detecting such intermediates.

The results of the off-line experiments in the comparison of mixtures of various porphyrins with tryptophan showed that the reaction products are the same for the three porphyrins studied (HP, TSO$_3$PP, and TMePyP). The TMePyP was run at both pH 10.4 and pH 7 in order to determine if the biologically known oxidation pathway to kynurenine could be induced. The results suggest that even at the physiological pH of 7, there is sufficient ammonia present in the buffer solution to induce the formation of the quinazoline derivatives.
CHAPTER 5
CONCLUSIONS AND FUTURE WORK

Cycloalkanoporphyrins

The analysis of 22 cycloalkanoporphyrin standards with six different skeletal types by tandem mass spectrometry (MS/MS) indicates that the skeletal type can be distinguished by carefully examining the daughter ion spectra of the compounds. When the normalized intensities of three ions in the daughter spectra ([M-43]+, [M-44]+, and [M-45]+) are plotted on a ternary diagram, distinct clusters of points are formed for each of the skeletal types studied.

The size of the isocyclic ring can be ascertained by the following diagnostic features in the daughter ion spectra. The CAP-5 and CAP-5-Me porphyrins display a dominant [M-44]+ ion in the [M-43]+ to [M-45]+ region of the spectrum. The CAP-5 and CAP-5-Me porphyrins are not distinguished by EIMS/MS. The CAP-6 porphyrins are characterized by a dominant [M-45]+ ion in this region of the mass spectrum. The presence of [M-45]+ and [M-44]+ ions in similar relative abundance is indicative of CAP-6-Me porphyrins. The CAP-7 porphyrins show the following pattern of daughter ions: [M-43]+ > [M-45]+ > [M-44]+. THB porphyrins are characterized by a dominant [M-43]+ ion.
Areas of future interest with this project would include a re-evaluation of the use of porphyrins in fingerprinting for oil-oil and oil-source rock correlations. It would be interesting to better characterize geoporphyrin extracts using this technique. The mechanisms for the formation of the \([M-43]^+, [M-44]^+,\) and \([M-45]^+\) ions are unclear. It would be of fundamental interest to determine the mechanisms of the fragmentations by isotopic labelling studies. The best place to start in this study would be to label the ring carbons and/or the carbons of the alkyl groups adjacent to the isocyclic ring. In light of the results from Chapter 3 of this dissertation, it would be interesting to determine how much of an effect the steric repulsions of the ring have on the anomalous fragmentations of the CAPs. Another area of interest would be to study the effects of extended alkyl chains (high carbon-number porphyrins) on the fragmentations of the CAPs.

**5-NO₂ OEP and Its Divalent Metal Complexes**

The fragmentation of 5-NO₂ OEP was found to be significantly different from that of OEP, with loss of OH⁻ and α-cleavage becoming important modes of fragmentation. MS/MS and HRMS data provide evidence for pyrrole ring scission in 5-NO₂ OEP. The loss of C₃H₅O⁺ from the molecular ion (m/z 579) forms an ion at m/z 522 that also appears in the parent ion spectrum of the m/z 375 ring-scission fragment. Previously, ring cleavage of non-substituted porphyrins has been observed only as a surface-induced phenomenon in CIMS. The ring-
cleavage process has been attributed to the non-planarity and subsequent loss of aromiticity exhibited by the 5-NO\textsubscript{2} OEP.

With the insertion of a divalent metal (Mg(II), Ni(II), Cu(II), Co(II), Zn(II), Ag(II), or Pd(II)) into 5-NO\textsubscript{2} OEP, the modes of fragmentation become significantly altered. A peak at either [M-75]\textsuperscript{+} or [M-86]\textsuperscript{+} was the base peak in the daughter spectra of the M\textsuperscript{+} ions of all of the metallated species. The [M-75]\textsuperscript{+} was attributed to loss of NO\textsubscript{2} and C\textsubscript{2}H\textsubscript{5}\textsuperscript{+}. The [M-86]\textsuperscript{+} peak involved oxygen migration in a process similar to that described for the non-metallated 5-NO\textsubscript{2} OEP. The fragmentation involves the loss of C\textsubscript{3}H\textsubscript{5}O\textsuperscript{+} analogous to 5-NO\textsubscript{2} OEP with the loss of an additional C\textsubscript{2}H\textsubscript{5}\textsuperscript{+}. The ring-cleavage phenomenon was not observed for any of the metal complexes studied. This effect can be attributed to the stabilization incurred by templating of the metal ion with the macrocycle.

The intensity of the ions in the daughter spectra of the M\textsuperscript{+} ions showed a marked dependence on the identity of the chelated metal ion. The [M-86]\textsuperscript{+} and the [M-15]\textsuperscript{+} (\(\beta\)-cleavage) ions were observed to be the most metal-dependent of the fragment ions. Plots of the 3050 cm\textsuperscript{-1} IR band and of the \(\nu_{10}\) Raman band versus the [M-86]\textsuperscript{+} daughter ion showed a similar trend. The 3050 cm\textsuperscript{-1} band corresponds to the C\textsubscript{meso}-H stretch, while the \(\nu_{10}\) Raman band corresponds to the C\textsubscript{\(\alpha\)}-C\textsubscript{meso}-C\textsubscript{\(\alpha\)} vibration. Both of these bands have been shown to be sensitive to the planar-nonplanar conformations of metallated porphyrins.

Future work in this area would include the investigation of other related compounds such as the polynitro OEPs, other meso-substituted porphyrins,
meso-nitro porphine, and tetraphenylporphyrin with a nitro group at the β-pyrollic position. Areas of fundamental interest would be to refine the investigations done here by incorporating data for other metallated OEPs. Also, as X-ray structural data and molecular modeling data become available, it would be interesting to correlate with the metal-dependency of the MS/MS studies.

Porphyrin Photosensitizers

The results of the on-line monitoring of porphyrin photosensitizing reactions showed that the flow cell design needs improvement in order to detect the short-lived intermediates of the reactions. It may also be necessary to change the type of ionization interface used, since porphyrin compounds were not detected by thermospray. On-line electrospray with laser illumination of the photosensitizer could provide a feasible way to detect intermediates. It may even be possible to illuminate directly through the spray region and photolyze biological compounds during the ionization process. The flow rates of electrospray are low enough that it would also be possible to use a capillary as the photolysis cell. Also, electrospray would not have the problem of the high backpressures associated with thermospray.
REFERENCES


Johnson, J.V.; Yost, R.A. Anal. Chem. 1985, 57, 758A.


Ortiz de Montellano, P.R. Cytochrome P450 Structure, Mechanism and Biochemistry, Plenum Press, New York, NY 1986.


Volk, K.J.; Yost, R.A.; Brajter-Toth, A. *Anal. Chem.* 1992, 64, 21A.


BIOGRAPHICAL SKETCH

John D. Laycock was born in Silver Springs, Maryland on January 30th, during the twenty-five year blizzard of 1966. His family moved to Pensacola, Florida when he was four and has lived there ever since. John graduated from Woodham High School in 1984; as an All-state placekicker he was offered an athletic scholarship to play football at Nicholls State University (Thibodeaux, LA). During his first semester at Nicholls, not wanting to deal with punch cards anymore, he changed his chosen major from Computer Science to Chemistry. Later that year, he was awarded the Outstanding Achievement in Freshman Chemistry Award.

He graduated from Nicholls in May, 1989 and continued on to graduate school that fall at the University of Florida. He has been working under the supervision of Richard A. Yost and in close collaboration with J. Martin E. Quirke of Florida International University in the tandem mass spectrometric study of porphyrins. Upon completion of his degree, he will be working as an analytical chemist, preferably somewhere near a good surfing beach.
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Richard A. Yost, Chair
Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Willard W. Harrison
Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James D. Winefordner
Graduate Research Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William Weltner
Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Anthony F. Randazzo
Professor of Geology
This dissertation was submitted to the Graduate Faculty of the Department of Chemistry in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as a partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 1994

Dean, Graduate School
ON MODEL FITTING FOR MULTIVARIATE POLYOMOUS RESPONSE DATA

By

JOSEPH B. LANG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1992
ACKNOWLEDGMENTS

I would like to express my appreciation to Dr. Alan Agresti for serving as my dissertation advisor. For the many comments, ideas, and lessons he has shared with me, I am greatly indebted. Through his advisement and guidance, he has taught me to appreciate and respect good statistical research and teaching. He is a mentor worthy of emulation. I also want to express my gratitude to Dr. Jane Pendergast, who also served on my dissertation committee. I learned a great deal from her during the two years that I worked in the Biostatistics Department. To all of the faculty at the University of Florida, I extend my thanks. The statistics department, with its scholarly and friendly atmosphere, proved to be a wonderful place to learn.

The influences of persons from my past are not forgotten. Without Patrick Kearin’s stimulating teaching of high school math, I may never have become interested in this subject. The genuine excitement delivered by Dr. James Kepner, in his teaching of undergraduate statistics, was the reason I decided to pursue an advanced degree in statistics.

I would like to thank my parents and the rest of my family for all of the support and encouragement they have given over the course of my studies and research. My friends and student colleagues deserve many thanks as well. Finally, I would like to thank Kendra Paar for always being there to support and encourage me while I was writing this paper.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>CHAPTERS</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 A Brief Introduction to the Problem</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Outline of Existing Methodologies—No Missing Data</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Outline of Existing Methodologies—Missing Data</td>
<td>12</td>
</tr>
<tr>
<td>1.4 Format of Dissertation</td>
<td>14</td>
</tr>
<tr>
<td>2 RESTRICTED MAXIMUM LIKELIHOOD FOR A GENERAL CLASS OF MODELS FOR</td>
<td>17</td>
</tr>
<tr>
<td>POLY TOMOUS RESPONSE DATA</td>
<td></td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>17</td>
</tr>
<tr>
<td>2.2 Parametric Modeling—An Overview</td>
<td>24</td>
</tr>
<tr>
<td>2.2.1 Model Specification</td>
<td>25</td>
</tr>
<tr>
<td>2.2.2 Measuring Model Goodness of Fit</td>
<td>33</td>
</tr>
<tr>
<td>2.3 Multivariate Polytomous Response Model Fitting</td>
<td>43</td>
</tr>
<tr>
<td>2.3.1 A General Multinomial Response Model</td>
<td>44</td>
</tr>
<tr>
<td>2.3.2 Maximum Likelihood Estimation</td>
<td>48</td>
</tr>
<tr>
<td>2.3.3 Asymptotic Distribution of Product-Multinomial ML Estimator</td>
<td>56</td>
</tr>
<tr>
<td>2.3.4 Lagrange's Method—The Algorithm</td>
<td>60</td>
</tr>
<tr>
<td>2.4 Comparison of Product-Multinomial and Product-Poisson Estimators</td>
<td>67</td>
</tr>
<tr>
<td>2.5 Miscellaneous Results</td>
<td>78</td>
</tr>
<tr>
<td>2.6 Discussion</td>
<td>83</td>
</tr>
</tbody>
</table>
3 SIMULTANEOUSLY MODELING THE JOINT AND MARGINAL DISTRIBUTIONS OF MULTIVARIATE POLYTOMOUS RESPONSE VECTORS ................. 87

3.1 Introduction .................................................. 87
3.2 Product-Multinomial Sampling Model ....................... 88
3.3 Joint and Marginal Models .................................. 93
3.4 Numerical Examples .......................................... 98
3.5 Product-Multinomial Versus Product-Poisson Estimators: An Example .................. 111
3.6 Well-Defined Models and the Computation of Residual Degrees of Freedom ........ 121
3.7 Discussion ..................................................... 132

4 LOGLINEAR MODEL FITTING WITH INCOMPLETE DATA .......... 135

4.1 Introduction .................................................. 135
4.2 Review of the EM Algorithm .................................. 137
  4.2.1 General Results ........................................ 138
  4.2.2 Exponential Family Results ............................. 140
4.3 Loglinear Model Fitting with Incomplete Data ................. 144
  4.3.1 The EM Algorithm for Poisson Loglinear Models .... 145
  4.3.2 Obtaining the Observed Information Matrix .......... 148
  4.3.3 Inferences for Multinomial Loglinear Models ......... 152
4.4 Latent Class Model Fitting—An Application ................. 160
4.5 Modified EM/Newton-Raphson Algorithm .................... 166
4.6 Discussion ..................................................... 170

APPENDICES

A CALCULATIONS FOR CHAPTER 2 .............................. 172
B CALCULATIONS FOR CHAPTER 4 ............................... 176

BIBLIOGRAPHY .................................................... 193

BIOGRAPHICAL SKETCH ........................................... 200
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Opinion Poll Data Configuration</td>
<td>22</td>
</tr>
<tr>
<td>3.1</td>
<td>Interest in Political Campaigns</td>
<td>91</td>
</tr>
<tr>
<td>3.2</td>
<td>Cross-Over Data</td>
<td>92</td>
</tr>
<tr>
<td>3.3</td>
<td>Joint Distribution Models—Goodness of Fit</td>
<td>100</td>
</tr>
<tr>
<td>3.4</td>
<td>Marginal Distribution Models—Goodness of Fit</td>
<td>101</td>
</tr>
<tr>
<td>3.5</td>
<td>Candidate Models in $J(L \times L + D) \cap M(U)$—Goodness of Fit</td>
<td>102</td>
</tr>
<tr>
<td>3.6</td>
<td>Estimates of Freedom Parameters for Model $J(L \times L + D) \cap M(CU)$</td>
<td>103</td>
</tr>
<tr>
<td>3.7</td>
<td>Freedom Parameter Estimates and Standard Errors</td>
<td>105</td>
</tr>
<tr>
<td>3.8</td>
<td>Estimated Cell Means and Standard Errors</td>
<td>106</td>
</tr>
<tr>
<td>3.9</td>
<td>Cross-Over Data Models—Goodness of Fit</td>
<td>110</td>
</tr>
<tr>
<td>3.10</td>
<td>Freedom Parameter ML Estimates for Model $J(UA) \cap M(U)$</td>
<td>110</td>
</tr>
<tr>
<td>3.11</td>
<td>Children’s Respiratory Illness Data</td>
<td>112</td>
</tr>
<tr>
<td>3.12</td>
<td>Product-Multinomial versus Product-Poisson Freedom Parameter Estimation</td>
<td>117</td>
</tr>
<tr>
<td>4.1</td>
<td>Observed cross-classification of 216 respondents</td>
<td>162</td>
</tr>
<tr>
<td></td>
<td>with respect to whether the tend toward</td>
<td></td>
</tr>
<tr>
<td></td>
<td>universalistic (1) or particularistic (2) values</td>
<td></td>
</tr>
<tr>
<td></td>
<td>in four situations (A, B, C, D) of role conflict</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>Parameter and Standard Error Estimates</td>
<td>164</td>
</tr>
<tr>
<td>4.3</td>
<td>Classification Probability Estimates</td>
<td>165</td>
</tr>
</tbody>
</table>
Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

ON MODEL FITTING FOR MULTIVARIATE POLYTOPHOMOUS RESPONSE DATA

By

Joseph B. Lang

May, 1992

Chairman: Dr. Alan Agresti
Major Department: Statistics

A broad class of models that imply structure on both the joint and marginal distributions of multivariate categorical (ordinal or nominal) responses is introduced. These parsimonious models can be used to simultaneously describe the marginal distributions of the responses and the association structure among the responses. As a special case, this class of models includes classical log- and logit-linear models. In this sense, we address model fitting for multivariate polytomous response data from a very general perspective. Simultaneous models for joint and marginal distributions are useful in a variety of applications, including longitudinal studies and studies dealing with social mobility and inter-rater agreement.

We outline a maximum likelihood fitting algorithm that can be used for fitting a large class of models that includes the class of simultaneous models. The algorithm uses Lagrange's method of undetermined multipliers and a modified Newton-Raphson iterative scheme. We also discuss goodness-of-fit tests and model-based inferences. Inferences for certain model parameters are shown to be equivalent for product-Poisson and product-multinomial
sampling assumptions. This useful equivalence result generalizes existing results. The models and fitting method are illustrated for several applications.

Missing data are often a problem for multivariate response data. We consider inferences about loglinear models for which only certain disjoint sums of the data are observable. We derive an explicit formula for the observed information matrix associated with the loglinear parameters that is intuitively appealing and simple to evaluate. The observed information matrix can be evaluated at the maximum likelihood estimates and inverted to obtain an estimate of the precision of the loglinear parameter estimates. The EM-algorithm can be used to fit these incomplete data loglinear models. We describe this algorithm in some detail, paying special attention to the Poisson loglinear model fitting case. Alternative fitting algorithms are also outlined. One proposed alternative uses both the EM and Newton-Raphson algorithm, thereby resulting in a faster, more stable, algorithm. We illustrate the utility of these results using latent class model fitting.
CHAPTER 1
INTRODUCTION

1.1 A Brief Introduction to the Problem

There are many situations when multiple responses are observed for each 'subject' in a group, or several groups. Here 'subject' is generically used to refer to a randomly chosen object that generates responses. The multiple responses could represent repeated measurements taken on subjects over time or occasions. They could be the ratings assigned by several judges that all viewed and rated the same set of slides (here, the 'subjects' are the slides). Or, perhaps, it may be that several distinct or noncommensurate responses are recorded for each subject. These responses are often categorical—ordinal or nominal—and inevitably interrelated. This dissertation addresses issues related to modeling and model fitting for multivariate categorical (ordinal or nominal) responses.

Models for multivariate categorical response data are usually developed to answer questions about (i) the association structure among the multiple responses or (ii) the behavior of the marginal distributions of the response variables. Specifically, a typical question of the first type is, "How are the responses interrelated and is this interrelationship the same across the levels of the covariates?" A typical type ii question is, "How do the (marginal) responses depend on the covariates or occasions?" Historically, many models (e.g. log- and logit-linear models) have been developed for the primary
purpose of answering the type i questions. Many of these models can easily be fitted using maximum likelihood (ML) methods. These models typically, however, are not useful for answering the type ii questions (Cox, 1972). Marginal models—those models used to answer type ii questions—are not as well developed. One reason for this is that ML fitting of these marginal models is more difficult. At present, the method of weighted least squares (WLS) is used almost exclusively for fitting these models.

Suppose that we are interested in answering questions of both types i and ii. Usually the questions are addressed using two different models, a joint distribution model and a marginal model, and fitting them separately. It seems reasonable to want a model that can be used to address simultaneously both questions. That is, we would like a model that simultaneously implies structure on both the joint and marginal distribution parameters. To date, there has been very little work done on the development and fitting of these simultaneous models.

Whenever multiple responses are observed it is inevitable that there will be missing data. There are several ways to fit the Poisson loglinear model with incomplete data. One popular method is to use the EM algorithm to find the ML estimates of the loglinear parameters. One drawback to this algorithm is that a precision estimate of the ML estimators is not produced as a by-product. Several numerical techniques have been developed to approximate the observed information matrix, which, upon inversion, will act as the precision estimate. However, it would be of some convenience to derive an explicit formula for the observed information matrix, at least in some special cases.
1.2 Outline of Existing Methodologies—No Missing Data

We begin our discussion by considering the case of no missing data. There are many methods for analyzing multivariate categorical (ordinal or nominal) response data. These methods usually involve fitting (separately) models for the joint or the marginal distributions of the response vectors. In rare instances, simultaneous models for both the joint and marginal distributions are considered. Maximum likelihood fitting methods for the joint distribution models are simple and described in almost every standard text on categorical data analysis. The fitting of marginal models using ML methods is more difficult. Maximum likelihood fitting of the marginal homogeneity model was considered by Madansky (1963) and Lipsitz (1988). The fitting of a more general class of marginal models was considered by Haber (1985a). Finally, the fitting of simultaneous models using ML methods has only been addressed in the bivariate response case. The fitting technique becomes very complicated when there are more than two categorical responses. To appreciate the complexity of extending the technique to multivariate response data, see section 6.5 of McCullagh and Nelder (1989) or perhaps Dale (1986). In contrast, the ML fitting method of Chapter 2 can easily be used to fit many marginal and simultaneous models. In the next few paragraphs, we briefly describe the existing methods for modeling and model fitting for multivariate categorical response data.

Modeling Joint Distributions Separately. One common method for analyzing multivariate categorical responses is to model the joint distribution only. These models, which include classical log- and logit-linear models for the
joint probabilities, are useful for describing the association structure among the responses. The last 30 years have seen the development of these methods for analyzing multivariate categorical responses (Haberman, 1979; Bishop et al., 1975; Agresti, 1984, 1990). For specificity, consider the following panel study: One hundred randomly selected subjects were asked how interested they were in the political campaigns. They were to respond on the 3 point ordinal scale, (1) Not Much, (2) Somewhat, and (3) Very Much. Then four years later the same group of subjects was asked to respond on the same scale to the same question. A separate investigation into the association structure would enable us to answer questions of a conditional nature. For example, we could estimate the probability of responding ‘Very Much’ on the second occasion given that the response at the first occasion was ‘Not Much’. The description of these ‘transitional’ probabilities, although very interesting, may not be completely satisfactory. We may also be interested in addressing questions with regard to the marginal distributions. Perhaps we would like to answer the question, “Are the distributions of responses to the political interest question the same for each occasion?” Laird (1991), in a nice review of likelihood-based methods for longitudinal analysis, mentions that the utility of classical log- and logit-linear models is restricted to two situations: (1) modeling the dependence of a univariate response on a set of covariates and (2) modeling the association structure between a set of multivariate responses. These models place structure on the joint probabilities and so they are not directly useful for studying the dependence of the marginal probabilities on occasion and other covariates. This problem was pointed out by several authors (Cox, 1972; Prentice, 1988; McCullagh and Nelder, 1989;
Liang et al., 1991). An advantage of these models is that they are simple to fit using either WLS (Grizzle et al., 1969), ML (McCullagh and Nelder, 1989), or iterative proportional fitting (Bishop et al., 1975) methods. There are many standard statistical programs available for fitting these models (SAS, SPSS*, BMDP, GLIM, GENSTAT).

Modeling Marginal Distributions Separately. A second approach to analyzing multivariate categorical responses is to model only the marginal distributions and to ignore the joint distribution structure. Full likelihood methods that consider only models for the marginal probabilities tacitly assume a saturated model for the joint distribution. Therefore, the models may be far from parsimonious. In the non-Gaussian response setting, there is a distinction between these marginal models and the transitional (or conditional) models of the previous paragraph. Marginal models describe the occasion-specific distributions and the dependence of those distributions on the covariates. Transitional or conditional models describe the distribution of individual changes over occasions. Models for these transitions can be represented as probability distributions for the future state 'given' the past states. Questions regarding transition probabilities can only be investigated with longitudinal data. On the other hand, questions regarding the marginal probabilities could theoretically be answered using cross-sectional data, provided the cohort (subject) effects were negligible. Panel studies resulting in longitudinal data result in more powerful tests for significance of within cluster factors, such as occasion effect. This follows because there is a reduced cohort effect; we are using the same panel of subjects at each occasion. For
further discussion about the distinction between marginal and transitional models, see Ware et al. (1988), Laird (1991), and Zeger (1988).

We will briefly discuss existing methods for making inferences about the marginal probabilities separately. We will group these methods into 5 categories: (1) nonmodel-based methods, (2) WLS methods, (3) ML methods, (4) Semi-parametric methods, and (5) other methods.

Nonmodel-based methods can be used to derive test statistics used for testing specific hypotheses regarding the marginal distributions. Examples include the Cochran-Mantel-Haenszel (1950, 1959) statistic which can be used for testing the hypothesis of marginal homogeneity (MH) (cf. White et al., 1982), McNemar's (1947) statistic which can be used for testing the equality of two dependent proportions, and Madansky's (1963) likelihood-ratio statistic for MH. Madansky's statistic is a difference in fit of the model of marginal homogeneity to the fit of the unstructured (saturated) model (see also Lipsitz, 1988 and Lipsitz et al., 1990). Many other relevant test statistics, some of which are generalizations or modifications of the aforementioned (cf. Mantel, 1963; White et al., 1982), exist. Cochran's (1950) Q statistic and Darroch's (1981) Wald-type statistic are examples of other test statistics that can be used to test for marginal homogeneity.

Presently, if one was to fit a marginal model, say a generalized loglinear model of the form \( C \log A \mu = X \beta \), where \( \mu \) is the vector of expected counts in the full contingency table, he or she would most likely use the WLS fitting algorithm. Most statistical software that fits these generalized loglinear models does so using WLS. There are some advantages to using WLS. It is computationally simple. Second-order marginal information is all that is
needed. And, the estimates are asymptotically equivalent to ML estimates. Some disadvantages are that covariates must be categorical, sampling zeroes create problems, and estimates are sensitive when second-order marginal counts are small. The WLS method for analyzing categorical data was originally outlined by Grizzle, Starmer and Koch (1969). Subsequently, marginal models for longitudinal categorical data, or more generally multivariate categorical response data, have been introduced and fitted using the WLS method (Koch et al., 1977; Landis and Koch, 1979; Landis et al., 1988; Agresti, 1989).

Maximum likelihood fitting of marginal models is more difficult since the model utilizes marginal probabilities, rather than joint probabilities to which the likelihood refers. When the responses are correlated, as they invariably are, the marginal counts do not follow a product-multinomial distribution. The full-table likelihood must be maximized subject to the constraint that the marginal probabilities satisfy the model. Haber (1985a) considers fitting generalized loglinear models of the form \( C \log A \mu = X \beta \) using Lagrange multipliers and an unmodified Newton-Raphson iterative scheme. The algorithm becomes very difficult to implement for even moderately large tables. This is primarily due to the difficulty of inverting the large Hessian matrix of the Lagrangian objective function. In this dissertation we consider a modified Newton-Raphson that uses a much simpler matrix than the Hessian. The matrix is easily inverted even for relatively large tables. Haber (1985b) considers the estimation of the parameters \( \beta \) in the special case \( C \log \mu = X \beta \). We will use a modification of the method of Aitchison and Silvey (1958, 1960) and Silvey (1959) to investigate the asymptotic behavior of the estimators of
\( \beta \) in the more general model \( C \log A\mu = X\beta \), thereby extending the work of Haber (1985b). Another relevant paper, Haber and Brown (1986), considers ML fitting of a model for the expected counts \( \mu \) that has loglinear and linear constraints. One can test hypotheses about the marginal probabilities by comparing the fit of relevant models. Haber (1985a, 1985b) and Haber and Brown (1986) only consider fitting the marginal models separately. No attempt has been made to simultaneously model the joint and marginal distributions.

Semi-parametric methods such as quasi-likelihood (Wedderburn, 1974) and a multivariate extension, generalized estimating equations (GEE), have become popular in recent years. The work of Liang and Zeger (1986), which advocated the use of these GEEs, has been extended to cover the multivariate categorical response data setting (Prentice, 1988; Zhao and Prentice, 1991; Stram et al., 1988; Liang et al., 1991). With these semi-parametric methods, the likelihood is not completely specified. Instead, generalized estimating equations are chosen so that, when the marginal model holds, even if the association among the multiple responses is misspecified, the estimators are consistent and asymptotically normally distributed. These estimators, used in conjunction with a robust estimator of their covariance (Liang and Zeger, 1986; Zeger and Liang, 1986; White, 1980, 1981, 1982; Royall, 1986), result in consistent inference about the effects of interest. When the responses are truly independent, the estimating equations with correlation matrix taken to be the identity matrix, are equivalent to the likelihood equations. The GEE approach requires the specification of a 'working' association or correlation matrix. Examples of working associations include those that imply all
pairwise associations (measured in terms of odds ratios) are the same and
that the higher order associations are negligible (Liang et al., 1991).

A related approach is known as GEE2. The consistency of these esti-
mators follows only if both the marginal model and the pairwise association
model are correctly specified. This approach is a second order extension
of the GEEs of Liang and Zeger (1986) which are now termed GEE1. It
is second order because the estimation of the marginal model parameters
and the pairwise association model parameters is considered simultaneously.
The focus of both approaches, GEE1 and GEE2, is usually on modeling
the marginal distributions—investigating how the marginal distributions
depend on occasion and covariates. The association is considered a nuisance.
Presently, there are no tests for goodness-of-fit of these models and so the
investigation into how well both models fit can be done only at an empirical
level. The assumption that higher order effects are negligible may not be
tenable. Testing procedures to assess the validity of these assumptions have
yet to be developed. Also, in contrast to WLS and ML methods, which
require only that the missing data be 'missing at random' (MAR), the semi-
parametric approaches require the missing data to be 'missing completely
at random' (MCAR). The assumption that the missing data mechanism is
MCAR is a much stronger assumption than MAR (Little and Rubin, 1986).

Finally, there are many other approaches to analyzing the marginal
probability structure separately. There are random effects models, whereby
subject-specific random effects induce a correlation structure on the multiple
responses. The marginal approach—the full likelihood is obtained by
averaging across the random effects—is computationally difficult (Stiratelli
et al., 1984). An alternative is to condition on the sufficient statistics for the subject effects and consider finding the estimates by maximizing the conditional likelihood. For further details on these conditional and unconditional methods see Rasch, 1961; Tjur, 1982; Agresti, 1991; Stiratelli et al., 1984; Conaway, 1989, 1990. As yet another alternative, Koch et al. (1980) give a bibliography for relevant nonparametric methods for analyzing repeated measures data. Agresti and Pendergast (1986) consider replacing the actual observations by their within cluster rank and testing for marginal homogeneity using the ordinary ANOVA statistic for repeated measures data. A three-stage estimator for repeated measures studies with possibly missing binary responses has been developed by Lipsitz et al. (1992). This approach is very similar to a generalized least squares approach, but it has some of the nice features of the GEE approaches. One of these nice features is that the estimators and their variance estimates are consistent under very mild assumptions. An extension of this method to the polytomous response case has yet to be developed.

Simultaneous Investigation of Joint and Marginal Distributions. There has been very little work done to investigate simultaneously the joint and marginal distribution structure. In some ways GEE2 is an attempt to describe both distributions. However, only the pairwise (not the joint) association structure is modeled; the higher-order associations are considered a nuisance. Tests comparing nested models have not been developed in this semi-parametric setting. Full likelihood approaches have been addressed by Dale (1986), McCullagh and Nelder (1989, Chapt. 6), and Becker and Balagtas (1991). Dale models the joint distributions of bivariate ordered
categorical responses by assuming that the log global odds ratios follow a linear model. The marginal probabilities are assumed to follow a cumulative logit model. McCullagh and Nelder consider simultaneously modeling the joint and marginal probabilities of a bivariate dichotomous response (two distinct responses) by assuming that the log odds-ratios follow a linear model and that the marginal probabilities follow a logit-linear model. Their example included age as a categorical covariate. Finally, Becker and Balagtas consider models for two-period cross-over data. The bivariate dichotomous response was the response to the two different treatments. Order of treatment application was considered a covariate. They assumed that the two log odds ratios followed a linear model and that the marginal probabilities satisfied a loglinear model. Because it is the marginal probabilities and not the joint probabilities that satisfy a loglinear model, Becker and Balagtas refer to the model as log nonlinear.

The ML model fitting approach used by each of these authors involves a reparameterization of the likelihood, which is a function of the joint probabilities, in terms of the joint and marginal model parameters. The reparameterization in the bivariate response case—the case each author considered—is somewhat complicated especially for multi-level responses. To make matters worse, the extension of this method to general multivariate polytomous responses looks to be extremely difficult. If the repaaparameterizations are made so that the full likelihood is expressible in terms of the joint and marginal model parameters, the likelihood can be maximized using a Newton-Raphson-type algorithm. Basically, one must solve for the root of some nonlinear score equation. This maximization approach is very sensitive
to the starting value in that convergence to a local maximum is not likely unless the starting estimate is very close to the actual maximum. Finding reasonable starting values is not a simple task. Dale (1986) outlines a method, specifically for the models considered in that paper, for finding a starting estimate.

In this dissertation, we outline an ML fitting method that can easily be used to fit a large class of simultaneous models, including those considered by Dale, McCullagh and Nelder, and Becker and Balagtas. The approach involves using Lagrange's method of undetermined multipliers along with a modified Newton-Raphson iterative scheme. For all of the models considered, an initial estimate for the algorithm is the data counts themselves along with a vector of zeroes corresponding to a first guess at the values of the Lagrange multipliers. The convergence of the algorithm is quite stable. The extension to multivariate polytomous response data is straightforward.

1.3 Outline of Existing Methodologies—Missing Data

Missing data is often an issue when the response is multivariate in nature. Missing data can also occur in more hypothetical situations. Examples include loglinear latent class models (Goodman, 1974; Haberman, 1988) and linear mixed or random effects models (Laird et al., 1987). In latent class analyses, a latent variable, which is unobservable, is assumed to exist. Mixed or random effects models posit the existence of some unobservable random variables that affect the mean response. In this brief outline, we will consider ML methods for model fitting when the data are not completely observable. Little and Rubin (1986) provide a nice summary of methods
for model fitting with incomplete data. There are many ways to find the maximum likelihood estimators when the data are not completely observable, each method having its positive and negative features. We could work directly with the incomplete-data likelihood, which is usually complicated relative to the complete-data likelihood, and use a Newton-Raphson or Fisher-scoring algorithm. Palmgren and Ekholm (1987) and Haberman (1988) use these methods to obtain maximum likelihood estimates and their standard errors. Alternatively, we could avoid the complicated likelihood altogether and use the Expectation-Maximization algorithm (Dempster et al., 1977). Sundberg (1976) discusses the properties of the EM algorithm when it is used to fit models to data coming from the regular exponential family. The EM algorithm is one of the more flexible ML fitting algorithms for missing data situations. We will primarily focus on this method for fitting loglinear models with incomplete data.

Although the EM algorithm is easily implemented to fit loglinear models with incomplete data, the algorithm does not provide an estimate of precision of the model parameter estimators. Meng and Rubin (1991) outline a supplemental EM (SEM) algorithm, whereby, upon convergence of the EM algorithm, the variance matrix for the model estimators is adjusted to account for missing data. The adjustment is a function of the rate of convergence of the EM algorithm, which in turn is a function of how much information is missing. Meng and Rubin numerically estimate the rate of convergence, thereby obtaining an estimate of precision that reflects missingness. Although this approach should prove to be applicable in the general situation, it still is desirable to derive an explicit formula for the variance matrix that reflects
missingness. Other authors (Meilijson, 1989; Louis, 1982) have discussed methods for estimating precision of model estimators when the data are incomplete and the EM algorithm is used. Meilijson's method involves EM-aided differentiation, which is essentially a numerical differentiation of the score vector. The method relies on the assumption that the observed data components are i.i.d. (identically and independently distributed). Louis gives an analytic formula for the observed information matrix based on the incomplete data. The computation of the observed information matrix based on this formula is not straightforward and must be considered separately for each special application.

1.4 Format of Dissertation

In Chapter 2, we develop a maximum likelihood method for fitting a large class of models for multivariate categorical response data. This development follows a general discussion about parametric modeling. Concepts such as degrees of freedom and model distances (or goodness of fit) are described at an intuitive level. We also describe and compare the asymptotic distributions of freedom parameter estimators under product-multinomial and product-Poisson sampling assumptions. Chapter 3 has more of an applied flavor. We consider simultaneously modeling the joint and marginal distributions of multivariate categorical response vectors. A broad class of simultaneous models is introduced. The models can be fitted using the techniques of Chapter 2. Several numerical examples are considered. Chapter 4 outlines the ML fitting technique known as the EM algorithm. This algorithm is used to fit models with incomplete data. Some advantages and disadvantages of using
the EM algorithm are addressed. The most important disadvantage is that the algorithm does not provide, as a by-product, a precision estimate of the ML estimators. We derive an explicit formula for the observed information matrix for the Poisson loglinear model parameters when only disjoint sums of the complete data are observable. An application to latent class modeling is considered. We also propose an ML fitting algorithm that uses both EM and Newton-Raphson steps. The modified algorithm should prove to have many positive features.

In this dissertation, we do not distinguish typographically between scalars, vectors, and matrices. Parameters and variables are treated as objects, their dimensions either being explicitly stated or implied contextually. By convention, functions that map scalars into scalars, when applied to vectors, will be defined componentwise. For example, if $\mu$ represents an $n \times 1$ vector, then

$$\log \mu \equiv (\log \mu_1, \log \mu_2, \ldots, \log \mu_n)' .$$

We frequently use abbreviations that are common in the statistical literature. They include ML (Maximum Likelihood), WLS (Weighted Least Squares), IWLS (Iterative (Re)Weighted Least Squares), and EM (Expectation-Maximization).

The range (or column) space of an $n \times p$ matrix $X$ is denoted by $\mathcal{M}(X)$ and is defined as $\{ \mu : \mu = X\beta, \beta \in \mathbb{R}^p \}$. The symbols $\otimes$ and $\oplus$ are the binary operators 'direct product' and 'direct sum'. The direct (or Kronecker) product is taken to be the right-hand product. That is,

$$A \otimes B = \{ Ab_{ij} \} .$$
The direct sum, $C$, of two matrices $A$ and $B$ is defined as

$$C = A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$ 

The symbol $D(\mu)$ represents a diagonal matrix with the elements of $\mu$ on the diagonal. That is,

$$D(\mu) = \begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n \end{pmatrix}.$$ 

In Chapter 4, we make use of the bracket notation often used by statistical and mathematical programming languages (e.g. Splus, Matlab). To illustrate the notation, consider a matrix $A$. The (sub)matrix $A[-2]$ is then matrix $A$ with the second column deleted. Similarly, the matrix $A[-3]$ is the matrix $A$ with the third row deleted.

Equation numbering is consecutive within sections of a chapter, the first number representing the chapter in which it appears. For example, the thirteenth equation in section 2.3 is equation (2.3.13). Within each appendix, the equations are numbered consecutively. For example, the third equation in Appendix B is numbered (B.3). Tables are numbered consecutively within chapters so that, for instance, Table 3.2 represents the second table within Chapter 3. Theorems, lemmas, and corollaries are numbered independently of each other. All are numbered consecutively within sections. Therefore, Corollary 3.2.2 is the second corollary within section 3.2 and Theorem 2.3.1 is the first theorem within section 2.3.
CHAPTER 2
RESTRICTED MAXIMUM LIKELIHOOD FOR A GENERAL
CLASS OF MODELS FOR POLYTOMOUS RESPONSE DATA

2.1 Introduction

In this chapter, we consider using maximum likelihood methods to fit a
general class of parametric models for univariate or multivariate polychotomous
response data. The models will be specified in terms of freedom equations
and/or constraint equations. These two ways of specifying models will be
discussed at length in section 2.2. The model specification equations may be
linear or nonlinear in the model parameters. Specifically, if \( \mu \) represents the
\( s \times 1 \) vector of expected cell means, the linear constraints will be of the form
\( L\mu = d \) and the nonlinear constraints will be of the form \( U'C\log(A\mu) = 0 \).
The freedom equations will have form \( C\log(A\mu) = X\beta \), where the
components of the vector \( \beta \) are referred to as the freedom parameters. In
Chapter 3 of this dissertation, we discuss more specifically models that can
be specified in terms of these constraint and freedom equations. The models
of that chapter allow one to simultaneously model the joint and marginal
distributions of multivariate polychotomous response vectors.

The maximum likelihood, model fitting algorithm of this chapter utilizes
Lagrange multipliers and a modified Newton-Raphson iterative scheme. In
particular, the models will be specified in terms of constraint equations and
the log likelihood will be maximized subject to the constraint equations being
satisfied. One common optimization algorithm found in the mathematics literature is Lagrange's method of undetermined multipliers. We show that Lagrange's method is easily implemented for ML fitting of the models under consideration in this chapter. One problem with Lagrange's method of undetermined multipliers for ML fitting of statistical models has been that it becomes computationally infeasible for large data sets. By using a modified Newton-Raphson method which involves inverting a matrix of a simpler form than the more complicated Hessian, we consider fitting models to relatively large data sets.

We also explore the asymptotic behavior of the estimators within the framework of constraint—rather than freedom—models. Usually, asymptotic properties of model and freedom parameter estimators are studied within the framework of freedom models. Aitchison and Silvey (1958, 1960) and Silvey (1959) studied the asymptotic behavior of the model parameter estimators when the model is specified in terms of constraint equations. Following the arguments of Aitchison and Silvey, we derive the asymptotic distributions of both the model and freedom parameter estimators.

Previous work by Haber (1985a) addressed maximum likelihood methods for fitting models of the form

\[ C \log(A\mu) = X\beta, \]

to categorical response data. Subsequently, Haber and Brown (1986) discussed ML fitting for loglinear models that were also subject to the linear constraints \( L\mu = d \), where these constraints necessarily include the identifiability constraint required of \( \mu \), the vector of product-multinomial
cell means. Both of these papers advocated the use of Lagrange's method of undetermined multipliers to find the maximum likelihood estimates of the model parameters \( \mu \). The method of Haber (1985a) involved using the (unmodified) Newton-Raphson method which becomes computationally unattractive as the number of components in \( \mu \) gets moderately large. Both Haber (1985a) and Haber and Brown (1986) were primarily concerned with measuring model goodness of fit and therefore did not consider estimation of freedom parameters. Haber (1985b) did consider estimation of freedom parameters, but only when the simpler model \( C \log \mu = X \beta \) was used. One of the several ways that we extend the work of Haber (1985a, 1985b) and Haber and Brown (1986) is to consider estimation of the freedom parameters when the more general model \( C \log A \mu = X \beta \) is used.

Others have considered ML fitting of nonstandard models for multivariate polytomous response data. Laird (1991) outlines the different approaches taken by different authors. As an example, Dale (1986) considered ML fitting for a particular class of models for bivariate polytomous ordered response data which were of the form

\[
C_1 \log(A_1 \mu) = X_1 \beta_1, \quad g(A_2 \mu) = X_2 \beta_2
\]

Specifically, the first freedom equation specifies a loglinear model for the association between the two responses measured by the global cross-ratios (cross-product ratios of quadrant probabilities) so that \( C_1 \) and \( A_1 \) are of a particular form. The second set of freedom equations specifies some generalized linear model (McCullagh and Nelder, 1989) for the marginal means or probabilities. Maximum likelihood estimators for the association