Characterization of integral membrane proteins using mass spectrometry : bacteriorhodopsin, rhodopsin, and neutrophil cytochrome b558
by David Robert Barnidge

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry
Montana State University

Abstract:
Electrospray ionization mass spectrometry (ESI MS) and matrix assisted laser desorption/ionization mass spectrometry (MALDI MS) have been used extensively to characterize the structural properties of cytosolic proteins. However, ESI and MALDI have had limited success in determining the accurate mass of integral membrane proteins, presumably due to interference from the detergents needed to solubilize them. Therefore, mass spectrometry has had limited applicability in structural studies of integral membrane proteins.

The integral membrane protein bacteriorhodopsin (BR) was used as a model in the development of a method for determining the accurate mass of BR solubilized in the detergents cetyltrimethylammonium bromide (CTAB), octyl-β-glucoside (OG), and CHAPSO by ESI MS. Quality mass spectra were obtained by means of a phase-separation technique where BR was partitioned into a chloroform-rich layer after mixing BR solubilized in an aqueous detergent solution with a chloroform:methanol:water mixture.

Tryptic digestion of the G protein-coupled receptor rhodopsin was used as a test case for mapping the transmembrane peptides of an integral membrane protein using MALDI MS. Affinity purified rhodopsin solubilized in OG was digested with trypsin and the resulting proteolytic fragments were separated using a modified reverse-phase high performance liquid chromatography technique. Detergent was incorporated throughout the procedure. Transmembrane tryptic peptides were identified by accurate mass and confirmed by Edman microsequencing.

The heterodimeric integral membrane protein cytochrome b558 from human neutrophils is the central component of the microbicidal NADPH-oxidase enzyme system. The protein shuttles electrons across the lipid bilayer via hemes located within the hydrophobic regions of the molecule to generate superoxide anion. Little is known about the tertiary structure of the protein, including the heme binding regions. Methods were developed for isolating fragments of purified cytochrome b558 that retained heme absorbance spectra after partial digestion with either trypsin or endoprotease Glu-C (V8) for analysis by SDS-PAGE, Western Blotting, and MALDI MS. Evidence from these analyses suggested that a region of the small subunit containing histidine 94 and regions of the large subunit of cytochrome b558 is involved in binding one of the hemes.

CHARACTERIZATION OF INTEGRAL MEMBRANE PROTEINS USING MASS SPECTROMETRY: BACTERIORHODOPSIN, RHODOPSIN, AND NEUTROPHIL CYTOCHROME b_{558}

by

David Robert Barnidge

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Chemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

July, 1998

APPROVAL

of a thesis submitted by

David Robert Barnidge

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

July 22, 1998
Date

Chairperson, Graduate Committee

Approved for the Major Department

July 30, 1998
Date

Head, Major Department

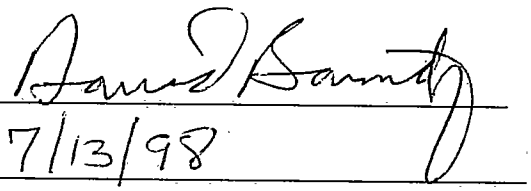
Approved for the College of Graduate Studies

August 4, 1998
Date

Joseph J. Fedock
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation for sale in and from microform or electronic format, along with the right to reproduce and distribute my abstract in any format in whole or in part."

Signature 

Date 7/13/98

This thesis is dedicated to my wife Phyllis who has given me the three most beautiful things in my life, Preston, Miles, and herself.

ACKNOWLEDGEMENTS

I would like to sincerely thank Dr. Jan Sunner and Dr. Al Jesaitis for their guidance and patience. Their willingness to allow me to combine the disciplines of mass spectrometry and protein chemistry is greatly appreciated. The experience I have gained from being exposed to these two diverse fields has made me a better analytical chemist. It has also given me the opportunity to explore the fascinating world of cellular biology. I would like to thank Dr. Ed Dratz for his helpful insights and encouragement. I would also like to thank Dr. Joe Sears for sharing his knowledge of mass spectrometry with me and for allowing me to become familiar with a variety of mass spectrometers.

I am grateful to all of my coworkers in the Jesaitis Laboratory, Drs. Heini Miettinen and John Mills, Kathy Billings-Sandoval, Heather Edens, Maria Kohler, and Michael Vlases. I would like to acknowledge the patience and support that was shown to me by Dan Siemsen when I showed up in the laboratory as a 'green' analytical chemist who didn't know a SDS-PAGE gel from a Western blot. Thanks Dan, for teaching me the best gradient gel technique around.

Most notably, I am indebted to Jeannie Gripentrog, Connie MacDonald, Tom Foubert, and Dr. Jim Burritt, the 'cytochrome crew', for all their help. To all my coworkers, thank you for making my tenure as a graduate student such an enjoyable, and often times hilarious, experience.

Finally I would like to thank my wife and my two children for their never-ending support.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

TABLE OF CONTENTS--Continued

TABLE OF CONTENTS--Continued

TABLE OF CONTENTS--Continued

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURE--Continued

LIST OF FIGURE--Continued

LIST OF FIGURE--Continued

LIST OF FIGURE--Continued

LIST OF FIGURE--Continued

LIST OF FIGURE--Continued

ABSTRACT

Electrospray ionization mass spectrometry (ESI MS) and matrix assisted laser desorption/ionization mass spectrometry (MALDI MS) have been used extensively to characterize the structural properties of cytosolic proteins. However, ESI and MALDI have had limited success in determining the accurate mass of integral membrane proteins, presumably due to interference from the detergents needed to solubilize them. Therefore, mass spectrometry has had limited applicability in structural studies of integral membrane proteins.

The integral membrane protein bacteriorhodopsin (BR) was used as a model in the development of a method for determining the accurate mass of BR solubilized in the detergents cetyltrimethylammonium bromide (CTAB), octyl-β-glucoside (OG), and CHAPSO by ESI MS. Quality mass spectra were obtained by means of a phase-separation technique where BR was partitioned into a chloroform-rich layer after mixing BR solubilized in an aqueous detergent solution with a chloroform:methanol:water mixture.

Tryptic digestion of the G protein-coupled receptor rhodopsin was used as a test case for mapping the transmembrane peptides of an integral membrane protein using MALDI MS. Affinity purified rhodopsin solubilized in OG was digested with trypsin and the resulting proteolytic fragments were separated using a modified reverse-phase high performance liquid chromatography technique. Detergent was incorporated throughout the procedure. Transmembrane tryptic peptides were identified by accurate mass and confirmed by Edman microsequencing.

The heterodimeric integral membrane protein cytochrome b_{558} from human neutrophils is the central component of the microbicidal NADPH-oxidase enzyme system. The protein shuttles electrons across the lipid bilayer via hemes located within the hydrophobic regions of the molecule to generate superoxide anion. Little is known about the tertiary structure of the protein, including the heme binding regions. Methods were developed for isolating fragments of purified cytochrome b_{558} that retained heme absorbance spectra after partial digestion with either trypsin or endoprotease Glu-C (V8) for analysis by SDS-PAGE, Western Blotting, and MALDI MS. Evidence from these analyses suggested that a region of the small subunit containing histidine 94 and regions of the large subunit of cytochrome b_{558} is involved in binding one of the hemes.

Chapter 1

INTRODUCTION

Electrospray Ionization and Matrix Assisted Laser Desorption/Ionization Mass Spectrometry

Over the past decade mass spectrometry has become established as an essential analytical technique for use in the characterization of proteins and peptides. This important development has resulted largely because electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) (1-7) that have been able to provide gas-phase ions for accurate mass determination of previously unionizable compounds. Fast atom bombardment, or FAB, was the first ionization technique that produced high mass (≤ 1,000 Da) gas-phase ions. In FAB, ions are generated from a viscous liquid, usually glycerol, by bombarding the surface with energetic atoms or ions. The generation of gas-phase ions from peptides having a mass above a few kDa was, however, unsuccessful (8). In spite of the this relatively low upper mass limit, FAB was used extensively for producing gas-phase ions from a variety of biopolymers thus starting a new field of

analytical chemistry termed 'bioanalytical' mass spectrometry.

It was not until the mid 1980's that the ionization techniques ESI and MALDI were introduced. In 1985, Fenn and coworkers in the Department of Chemical Engineering at Yale University devised a method to produce gas phase ions from a solution at atmospheric pressure for subsequent analysis by a quadrupole mass spectrometer (9). By applying a high voltage (~2,000 V) to the tip of a thin capillary, connected to a liquid introduction system, they were able to produce an 'electrospray' generating μm sized droplets that contained charged analyte molecules. The generation of a spray induced by a high electric field was not novel since others had reported the production of such 'electrosprays' much earlier, however, it was not known at the time that gas phase ions were produced. Yet, Fenn et al. were the first to couple an electrospray source at atmospheric pressure to a heated capillary backed by a multi-stage vacuum system. This advance allowed the introduction of gas-phase ions produced by the charged droplets into the mass spectrometer without overloading the high vacuum system.

The mechanism for the transfer of ions from solution into the gas-phase, derives from a competition between the Coulombic repulsion of like charges in a droplet and the surface tension of the droplet. The charge density on the surface of the droplet increases due to evaporation of neutral solvent ions until it reaches the Rayleigh limit where Coulombic repulsion becomes large enough to overcome the surface tension of the droplet. At this point the droplet becomes unstable and decomposes, producing daughter droplets. This process repeats itself until the droplets reach a small enough size such that the charge density on the surface induces a field strong enough to desorb the ion from the

droplet(10). This field induced desorption ia a 'soft' form of ionization that results in even-electron ions with low internal energies.

Frequently, clusters of analyte ions with neutral solvent molecules are produced. For accurate mass determination of the analyte, the ions must be completely desorbed. The elimination of clusters is accomplished either through the use of a heated ion transport capillary or heated counter current gas. Transporting the ions produced by the electrospray process into the high vacuum region of the mass spectrometer is achived through the use of lenses or small octopoles. Bare analyte ions are focused into the quadrupoles via ion optics then filtered by mass due to the action of the RF and DC fields applied to the quadrupole rods. A mass spectrum is produced by recording the signal from the detector as the RF and DC voltages on the quadrupoles are varied from lower to higher potential.

In the case of positively charged proteins and peptides, the extent of multiple charging observed in ESI is dependent upon the number of sites where protonation can occur. A mass spectrum of a protein electrosprayed from an acidic solution will reflect the number of solvent accessible basic sites. This has been used to probe conformational structures of proteins under different solvent conditions (11,12). Calculation of the accurate mass of the analyte is accomplished through the use of algorithms, however, the mass can easily be calculated manually. The presence of multiple peaks has the advantage of providing a more robust calculation of the accurate mass since the mass calculated for each peak is used to determine an average molecular weight. Multiple charging also has the advantage of producing ions with low m/z values. Thus, molecules with high

molecular weights (upwards of 100 kDa) may be observed using standard quadrupole mass analyzers that have a mass limit of approximately 3,000 Da for singly charged ions..

Though multiple charging in electrospray ionization has clear advantages in the determination of high molecular weight biopolymers, there are also some disadvantages. If for instance, a sample contains mutiple components, it is often difficult to interpret the spectrum, since overlapping of charge states from the different analytes can occur. Deconvoluting algorithms exist to aid in interpretation of complex spectra from heterogeneous samples. However, liquid chromatography or capillary electrophoresis is often used 'in-line' with the electrospray source to help avoid complex ES mass spectra. Since electrospray is a liquid phase atmospheric pressure ionization technique, it is superbly suited to the coupling of liquid phase separation methods to a mass spectrometer. Flow rates for routine LC (1-5 ml/min) are well above the flow rates used in electrospray (0.5-50 μl/min). Therefore, the LC flow is either split, or chromatography is performed using a 'micro' LC system. LC/ESI or CE/ESI coupled to a tandem mass spectrometer offers significant advantages. Tandem instruments are capable of determining the masses of daughter ions produced by collisional induced dissociation (CID) of electrospray generated pseudo-molecular ions. The analysis of daughter ions produced by CID is termed MS/MS. Mass spectrometers capable of performing MS/MS experiments include triple quadrupoles, ion traps, and FTICR mass spectrometers (13). The amino acid sequence of a parent or peptide ion can often be determined from a daughter ion MS/MS spectrum. MS/MS has been used extensively for sequencing of proteolytic peptides (2,5,14-21). This technique also enables sites of post-translational modifications such as

phosphorylation (22-24) and glycosylation (25-27) to be characterized using mass spectrometry.

Matrix-assisted laser desorption/ionization, or MALDI, is the second primary ionization technique used in bioanalytical mass spectrometry (28-31). In 1988 Karas and Hillenkamp at the University of Münster in Germany, found that ions are produced from a UV absorbing crystalline matrix when ablated with a UV laser in vacuum (32,33). This observation was, by itself, not surprising. However when analytes such as peptides and proteins were mixed with the matrix solution prior to crystallization, desorption/ionization of these analytes also occured. MALDI coupled to a time-of-flight mass spectrometer (TOF) is capable of detecting singly charged proteins with masses in excess of 100,000 Da. Singly charged ions of high molecular weight proteins can be observed since there is theoretically no upper mass limit for this type of mass analyzer (34).

The process of desorption/ionization in MALDI is not completely understood. However, it is recognized that the matrix used strongly affects the yield, mass resolution, and fragmentation of analyte ions (3,35,36). For each new analyte it becomes essential to try several different matrices. Some of the most commonly used matices are, dihydroxy benzoic acid (DHB), sinapininc acid (SA), and α-cyano-4-hydroxy cinnamic acid (α-CHCA). All of these matrices are small (<400 Da), conjugated organic acids having high extinction coefficients in the UV (37). Samples are prepared by combining a saturated solution of the matrix allowing the mixture to crystallize.

Many variations on this procedure have been reported in the literature. In most cases homogeneous or uniform crystal morphology upon drying results in more consistent

results. This is evident by the ability to generate analyte ions from different regions within the sample. Yet, even with optimum sample preparation, 'hot-spots' on the sample exist where analyte signal is more intense. In addition, suppression effects are seen for mixtures of analytes. An alternative method to MALDI has been described where the analyte is mixed with graphite or glycerol to produce what is known as SALDI or surface-assisted laser desorption/ionization (38). This technique has the advantage of low background from matrix ions and shows promise for adaptation to liquid systems.

Despite the problems in sample preparation, MALDI, coupled with TOF, is still a powerful technique for the accurate mass determination of proteins and peptides. Since MALDI makes use of a solid matrix, samples are introduced into the source region of a TOF mass spectrometer either on a standard insertion probe through a vacuum interlock, or on a sample plate transported through a series of pumping chambers. Once in the source, the sample is positioned to align with the focus of the pulsed laser beam. Ion extraction is obtained by applying a continuos high electric field over the sample (~30 kV/cm). Ions produced by the laser pulse that have the same polarity as the potential applied to the sample will pass through the flight tube to the detector. A photodiode provides the trigger for the digital oscilloscope that records the analog signal from the detector. The digital oscilloscope converts the analog signal to a digital signal and stores the spectra that can be downloaded to a personal computer for data manipulation. The time-of-flight for a MALDI ion is related to its mass by the formula shown below.

$$t = L \sqrt{\frac{m}{2 \cdot V_{acc} \cdot q}}$$

where t is flight time, L is the length of the flight tube, m is the mass of the ion, V_{acc} is the value of the accelerating voltage, and q is the charge of the ion. The mass resolution in a linear TOF mass spectrometer is typically a few hundred. Thus an ion having a mass of 300 Da could be resolved from an ion having a mass of 301 Da (i.e. $m/\Delta m$ = resolution). A number of different modifications have been devised for increasing the mass resolution in MALDI TOF instruments. One of the most common modification is the introduction of a reflector. The reflector, or ion mirror, serves to focus ions of different kinetic energies by reflecting the ions in a potential gradient. Ions with higher kinetic energies will penetrate deeper into the gradient field and thus travel a longer distance, arriving at the detector at the same time as ions with lower kinetic energies traveling a lesser distance. A major breakthrough in the pursuit of increased mass resolution for TOF instruments is the incorporation of delayed ion extraction technology (39). Instead of continuously extracting ions, the application of the extraction field is delayed a few μs after the laser pulse. The result is that the kinetic energies of the ions in the plume are more uniform due to collisionals within the plume. This uniformity translates into mass resolution on the order of several thousands in a linear TOF.

Comparison of ESI MS and MALDI MS of Hydrophobic Proteins and Peptides

ESI and MALDI have their respective strengths and weaknesses that derive from a number of key differences between the two techniques. Generally, MALDI is more

tolerable of buffers, salts, and detergents, used in the purification of proteins and peptides, than ESI (40). MALDI also uses much less material than standard ESI to produce mass spectra of comparable quality. Sample preparation in MALDI is quite simple, and the sample stages used in many instruments have multiple sites, therefore screening of a large number of samples is less time consuming using MALDI as compared to ESI. However, the mass accuracy and resolution of ESI quadrupole instruments is superior to that of linear MALDI TOF instruments without delayed ion extraction, especially for analytes with mass above 2,000 Da. To summarize, the optimum ionization technique is dependent upon the amount of sample available and its purity.

Both ESI and MALDI have been used extensively for the characterization of protein digests (3,41-48). In addition, a large number of reports have described accurate mass determinations of intact proteins using both ESI and MALDI (2,4,14). Yet, out of thousands of publications describing mass spectral analysis of proteins and peptides, very few (roughly less than 50) involve the mass analysis of hydrophobic samples. Even fewer articles deal with the mass analysis of integral membrane proteins. The lack of mass spectral results for this class of biopolymers is due to the problems associated with sample handling and ionization. In the case of ESI, it is found that ions must first be solubilized in solvents that can be electrosprayed before they can be desorbed. Therefore, hydrophobic proteins or peptides solubilized in aqueous buffers containing high levels of salt and detergents (>50 mM) must be solubilized in non-polar solvents prior to ESI. This requires the exchange of detergent solubilized proteins and peptides into non-polar solutions without aggregating the sample. MALDI on the other hand is a solid-phase technique and

integrating the sample solubilized in detergent into the matrix is experimentally easier to accomplish. Thus, MALDI is prefered over ESI since removal of detergent is not a prerequisite for ionization.

Intact Membrane Proteins Analyzed by ESI MS and MALDI MS

Schindler et al. first desribed methods dedicated to the ESI MS analysis of hydrophobic proteins and peptides (49). These authors found that mixed polar/non-polar solvent systems, such as chloroform:methanol:water, were able to keep hydrophobic proteins and peptides solubilized and were ionizable by ESI. The report included the first published ESI mass spectrum of the integral membrane protein bacteriorhodopsin (m.w. 26,784 Da). The protein was solubilized from purple membrane using hexafluoroisopropanol and/or neat formic acid. However, no detergents were used for solubilization in this study. The first work to demonstrate the ESI analysis of bacteriorhodopsin, solubilized in detergent, was published by Le Maire et al. (50). In this case, the protein was isolated on an SDS-PAGE gel then electroeluted into solution with subsequent removal of the detergent before analysis by ESI MS. Recently, a method was described by Hufnagel et al. (51) for the isolation of bacteriorhodopsin from purple membrane using a series of non-polar solvent washes to remove excess lipid. These authors analyzed the protein by ESI MS in the chloroform:methanol:water solvent system described by Schindler et al. (49). In this study bacteriorhodopsin solubilized in detergent prior to the solvent wash was also analyzed successfully by ESI MS.

MALDI MS has also been used to analyze the integral membrane protein

bacteriorhodopsin. The first report of an accurate mass for this protein, determined by MALDI MS, was by Beavis and Chait in a paper on the applicability of MALDI MS to the accurate mass determination of proteins (30). The work by Schey et al. demonstrated that MALDI MS could be used to determine the accurate mass not only of bacteriorhodopsin, but also of the G protein-coupled receptor rhodopsin (52). They also demonstrated that accurate masses could be obtained for the large peptides generated when rhodopsin was proteolyzed in rod outer segment membranes.

The samples used by Schey et al. were derived from non-detergent solubilized sources and this represents an unusual situation relative to most membrane proteins. Although it is promising that a native membrane protein can be ionized after a brief treatment with a non-polar solvent the fact remains that bacteriorhodopsin and rhodopsin are exceptions, in the sense that most mono-disperse preparations of membrane proteins are in detergent. Since most membrane proteins must be purified prior to mass spectral analysis, any methods used to determine the accurate mass of a protein must be able to tolerate the presence of detergent. The most thorough examination on the effect of different detergents on MALDI MS spectra of membrane proteins is the work by Rosinke et al. (53). The application of different detergents were examined using the membrane proteins bacteriorhodopsin, porin and cholesterolesterase. In addition, various matrixes were evaluated with the best obtained using dihydroxybenzoic acid. The non-ionic detergents such as octylglucoside (OG) were found to degrade the MALDI spectrum to a lesser extent than ionic detergents, such as sodium dodecyl sulfate (SDS).

Proteolytic Digests of Membrane Proteins Analyzed by ESI and MALDI

A majority of the peptide mapping experiments involving integral membrane proteins have focused on the hydrophilic fragments generated and not on the hydrophobic transmembrane fragments. Much of the mapping performed on integral membrane proteins, such as the acetylcholine receptor, were performed prior to the emergence of ESI and MALDI and instead relied on FAB for ionization of the proteolytic fragments (54,55). Also, early work on the FAB analysis of peptides produced by cyanogen bromide and chymotryptic digestion of both bacteriorhodopsin and rhodopsin were performed without any prior separation of the peptides (56). This deficiency was primarily due in part to the formidable difficulties involved in separating transmembrane peptides.

Attempts to overcome this problem have been made with limited success. MALDI MS was used in a study by Machold to identify a hydrophobic tryptic fragment of the nicotinic acetylcholine receptor that had been labeled with a radioactive photo-crosslinker specific for the α-neurotoxin binding site (57). The receptor was isolated by SDS-PAGE after reacting with the ligand and digested in the gel with trypsin. Small non-transmembrane fragments were extracted from the gel and isolated using conventional RP-HPLC. A similar technique was used in the study of sarcoplasmic reticulum Ca^{+2} ATPase by le Maire et al. (50) However, in this case the protein was digested with V8 while in vesicles, and SDS-PAGE was used to isolate the fragments that remained in the membrane after digestion. These authors isolation of these relatively large (>9,000 Da) hydrophobic peptides by electro-transferring the fragments into solution from select bands cut from the SDS-PAGE gel instead of isolating the fragments via RP-HPLC.

Nonetheless, the fragments, after electroelution from the gel into solution, were treated for removal of SDS prior to analysis by ESI.

Goals of This Dissertation: Advancement of Electrospray Ionization and Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Techniques for the Analysis of Hydrophobic Proteins and Peptides

It is clear from the preceeding discussion that there is a need for the development of methods designed to facilitate the accurate mass analysis of integral membrane proteins and their proteolytic peptides by ESI MS and MALDI MS. Montana State Unversity is fortunate enough to have both types of ionization techniques available in the form of a 'home-built' ESI source coupled to a VG TRIO-2 quadrupole mass spectrometer (58,59) and a Perseptive Biosystems Voyager MALDI time-of-flight mass spectrometer. The characterization of integral membrane proteins is at the forefront of research for a number of groups in the Department of Chemistry and Biochemistry, and the Department of Microbiology. Thus experiments were designed to address three questions regarding transmembrane proteins and their proteolytic fragments:

1) Can the accurate mass of the integral membrane protein bacteriorhodopsin be determined using electrospray ionization mass spectrometry, even in the presence of detergent?

2) Can transmembrane tryptic peptides from the detergent solubilized G protein-coupled receptor, rhodopsin, be isolated then analyzed using matrix-assisted laser

desorption/ionization mass spectrometry without the need for eliminating detergent?

3) Can these procedures be applied to transmembrane protein samples of lower purity?

The first part of this thesis is dedicated to the experiments that were designed to address the first two questions. The second part of the thesis focuses on the third question.

Human Phagocyte Cytochrome b_{558}

Cytochrome b_{558} from human phagocytes is a heme containing integral membrane protein found in secretory granules, specific granules, and plasma membranes of human neutrophils and macrophages (60-62). The protein functions as the central redox component for the multi-protein enzyme system known as NADPH oxidase (63). It is responsible for transferring electrons across the membrane to molecular oxygen to form superoxide, the high energy electron source for the generation of a number of microbicidal toxic oxygen compounds (64).

The increase in oxygen consumption by phagocytes exposed to opsonized bacteria is known as the 'respiratory burst'. When this 'respiratory burst' was first observed in 1933 it was thought to involve mitochondrial respiratory components (65). However, in 1959 it was shown that inhibitors of the reduction-oxidation centers in the mitochrondrial system, such as azide and cyanide, were ineffective at disrupting the consumption of oxygen by activated neutrophils (66). The cyanide-insensitive 'respiratory burst' has been

shown to be a critical component of the host defensive arsenal of vertebrates (67) and abscence of the 'respiratory burst' in humans inflicted with the genetic disorder chronic granulomatous disease, or CGD, results in chronic, recurrent, life-threatening infections for these individuals (68).

Further research found that the substrate for the enzyme involved in the reduction of oxygen to superoxide for production of the 'respiratory burst' was either NADH or NADPH (63). After this discovery, the enzyme involved in the reduction of molecular oxygen at the expense of NADPH was known as NADPH oxidase. The chemistry of this reaction is summerized by:

$$NADPH + 2\ O_2 \cdot \rightarrow NADP^+ + H^+ + 2\ O_2^{\cdot-}$$

Prior to the discovery of the NADPH oxidase system in neutrophils, a b-type cytochrome was found in the cell membranes and specific granules of these cells (69). Spectral studies of this protein showed that it had a room temperature Soret band at 414 nm in the oxidized form and a reduced minus oxidized absorbance spectrum with a Soret band at 426 nm and α- and β- bands at 558 nm and 528 nm, respectively (69). Data collected on this protein suggested that this b-type cytochrome was part of the NADPH oxidase. For example, the protein had a low midpoint redox potential of -245 mV, lower than any other mammalian cytochrome b (70) and low enough to participate in the electron transport chain between NADPH and molecular oxygen. In addition, the kinetics of superoxide production mirrored the steady-state reduction kinetics of cytochrome b_{558} in neutrophil membranes (70). However, some of the most convincing data suggesting that the protein was indeed involved in the production of superoxide was the fact that its

spectral activity could not be found in patients having CGD (71). The connection between the genetic disorder CGD and the protein product cytochrome b_{558} was finally proven when cDNA was cloned and sequenced (72) and antisera made against fusion protein of the gene product were shown to recognize one of the subunits of the purified protein from neutrophils (73). Thus cytochrome b_{558} from the membranes of human neutrophils was found to be the obligate electron transferase involved in shuttling high energy electrons from NADPH to molecular oxygen to form superoxide, and the abscence of cytochrome b_{558} leads to CGD.

Characteristics of Human Cytochrome b_{558}

Cytochrome b_{558} is a heterodimeric integral membrane protein found in neutrophils at approximately 7 nmol/cell. To purify the protein from membranes, detergents such as Triton X-100 and octylglucoside must be used to produce a monomeric distribution of protein. Also, prudent use of protease inhibitors are required during purification to avoid degradation of the protein by the abundant endogenous proteases found in neutrophils. Isolation of membranes involves the collection of fresh whole blood from normal human donors followed by separation of the neutrophils. These cells are then disrupted and membranes are isolated by centrifugation.

Parkos et al. were the first to purify cytochrome b_{558} from neutrophil membranes (74). The authors demonstrated that the protein is a heterodimer consisting of a glycosylated large subunit (a.k.a. **gp**91-*phox* referring to **glyco**protein 91 kDa m.w.-*ph*agocyte *ox*idase) and a non-glycosylated small subunit (a.k.a. p22-*phox*) (74). Huang et

al. later confirmed that the stoichiometry of the subunits is 1:1 (75). When the purified protein is observed by SDS-PAGE, the large subunit is seen as a broad or 'fuzzy' band between 72-110 kDa while the small subunit runs as a narrow sharp band at 22 kDa (74,76,77). Sequencing of the genes for each subunit indicates that the large subunit is comprised of 570 amino acids with a total molecular weight of 65,291 Da, while the small subunit is comprised of 195 amino acids with a total molecular weight of 20,957 Da (78). When the protein is deglycosylated by treatment with *N*-glycosidase F and analyzed using SDS-PAGE, the large subunit shifts from the broad band at M_r= 72-110kDa to a single band at approximately M_r= 57 kDa while the band for the small subunit does not change. When the protein is treated with 0.1 N NaOH to remove any O-linked oligosaccharide, no change in the mass of the large subunit, nor the small subunit, is observed. Therefore it has been inferred the large subunit contains only N-linked sites of glycosylation at asparigine residues and that neither the large or the small subunit contains O-linked sites of glycosylation at serine or threonine residues (74,79). Recently, Wallach et al. have advocated that the sites of glycosylation on the large subunit are asparagines 131, 148, and 239 (80). However, this hypothesis was formed from results obtained using incompletely processed large subunit expressed in canine membranes, not whole human cells. Thus the location of the glycosylation sites in the purified protein has yet to be determined.

The reduced form of the protein displays weak CO binding as evident by the decrease in the absorbance of the Soret band and the α and β bands in the presence of CO (70,81). The absorbance spectrum of the protein is sensitive to excess pyridine, and is

completely lost by the addition of 1 M NaCl, or denaturing reagents, thus indicating that the hemes are non-covalently bound (74,81). Extinction coefficients for the oxidized and reduced bands of cytochrome b_{558} have been calculated using the pyridine hemochrome assay (81,82). Values for the reduced minus oxidized $\Delta\epsilon$@ 426 nm range from 106-161 mM^{-1} cm^{-1} while the values for $\Delta\epsilon$@ 558 nm range from 21.7-29.3 mM^{-1} cm^{-1}.

The specific heme content of purified cytochrome b_{558}, calculated using the extinction coefficients quoted above, is 20-30 nmol/mg of protein (74). The theoretical value for a 1:1 protein to heme ratio would result in only 8.9 nmol of heme per milligram of protein. This observation suggests that there is more than one heme in the molecule. Optical absorbance studies on cytochrome b_{558} done *in situ* by Iizuka et al., found that at 77 K the α band at 558 nm splits into two distinct peaks when cells are stimulated under reducing conditions (83). Work by Cross et al. showed that there are indeed two non-identical hemes present in the molecule as evident from the deviations in the redox titration curves of normal and mutated forms of cytochrome b_{558}. The mid-point potentials for the two hemes were determined to be at -225 and -265 mV (84).

Topology of Cytochrome b_{558}

No X-ray crystal structure for cytochrome b_{558} exists at this time due to difficulties in purifying large amounts of the protein. Additionally, since it is an integral membrane protein it must be kept in detergent to keep it from aggregating and thus is most likely difficult to crystallize. Consequently, only low-resolution information on the orientation of the protein in the membrane is available. Hydropathy analysis of the cytochrome b_{558}

suggests that there are five, possibly six, hydrophobic transmembrane spanning regions on gp91-*phox* and three transmembrane regions on p22-*phox*. Extracellular regions on the large subunit include ^{150}SYLNFARKRIKNPEGGLYL^{172} and ^{246}KISEWGKIKEC^{256} as determined by FACS analysis (85)(Barnidge, unpublished results). Intracellular regions on the large subunit include ^{86} STRVRRQL^{93}, ^{450}FEWFADLLQLL^{457}, ^{490} FAVHHDEEKDVITG^{503} , and ^{554}ESGPRGVHFIF^{464} as determined by the inhibition of the interaction between the cytosolic protein p47 *phox* with cytochrome b_{558} by peptides mimicing these regions (61,86). Small subunit intracellular regions include ^{148}KQPPSNPPRPPPAE^{161} and ^{175}AGGPPGGPQVNPIPVTDEVV^{194} that were also found to be binding regions of p47 *phox* (86-88). In addition, the region between ^{181}GGPQVNI^{188} was found to be intracellular by FACS analysis after identification of this region as the binding site for a small subunit monoclonal antibody (89). There also is evidence that suggests that this anti-small subunit antibody recognizes not only the region ^{181}GGPQVNI^{188} but also ^{27}ATAGRF^{32} as part of a 'split-epitope' (Burritt, unpublished results).

The amino acid sequence of the pg91-*phox* and p22-*phox* is listed on the next page taken from the Swiss-Prot protein database.

Table 1

Amino Acid Sequence of Large Subunit Human Cytochrome b_{558}

SEQUENCE amino acids - 569; molecular weight - 65,204 Da; SWISS-PROT #: P04839

[1]GNWAVNEGLS IFVILVWLGL NVFLFVWYYR VYDIPPKFFY TRKLLGSALA LARAPAACLN

[61]FNCMLILLPV CRNLLSFLRG SSACCSTRVR RQLDRNLTFH KMVAWMIALH SAIHTIAHLF

[121]NVEWCVNARV NNSDPYSVAL SELGDRQNES YLNFARKRIK NPEGGLYLAV TLLAGITGVV

[181]ITLCLILIIT SSTKTIRRSY FEVFWYTHHL FVIFFIGLAI HGAERIVRGQ TAESLAVHNI

[241]TVCEQKISEW GKIKECPIPQ FAGNPPMTWK WIVGPMFLYL CERLVRFWRS QQKVVITKVV

[301]THPFKTIELQ MKKKGFKMEV GQYIFVKCPK VSKLEWHPFT LTSAPEEDFF SIHIRIVGDW

[361]TEGLFNACGC DKQEFQDAWK LPKIAVDGPF GTASEDVFSY EVVMLVGAGI GVTPFASILK

[421]SVWYKYCNNA TNLKLKKIYF YWLCRDTHAF EWFADLLQLL ESQMQERNNA

[471]GFLSYNIYLT GWDESQANHF AVHHDEEKDV ITGLKQKTLY GRPNWDNEFK TIASQHPNTR

[531]IGVFLCGPEA LAETLSKQSI SNSESGPRGV HFIFNKENF

Potential Transmembrane Residues

TRANSMEM ^{8}L to ^{29}Y TRANSMEM ^{98}T to ^{121}N TRANSMEM ^{166}L to ^{190}T

TRANSMEM ^{204}F to ^{223}A TRANSMEM ^{265}N to ^{285}V

Potential FAD and NADPH Binding Domains

FAD ^{406}V to ^{447}T NADPH ^{503}G to ^{512}R and ^{531}I to ^{569}F

Know Mutations found in X-linked CGD Patients

VARIANT ^{100}H -> R (IN X-CGD); VARIANT ^{155}A -> T (IN X-CGD)

VARIANT ^{208}H -> Y (IN X-CGD); VARIANT ^{243}C -> S (IN X-CGD)

VARIANT ^{338}P -> H (IN X-CGD); VARIANT ^{388}G -> A (IN X-CGD)

VARIANT ^{414}P -> H (IN X-CGD); VARIANT ^{499}D -> G (IN X-CGD)

Table 2

Amino Acid Sequence of Small Subunit Human Cytochrome b_{558}

SEQUENCE amino acids - 194; molecular weight - 20,827 Da; SWISS-PROT#: P13498

[1]GQIEWAMWAN EQALASGLIL ITGGIVATAG RFTQWYFGAY SIVAGVFVCL LEYPRGKRKK

[61]GSTMERWGQK HMTAVVKLFG PFTRNYYVRA VLHLLLSVPA GFLLATILGT ACLAIASGIY

[121]LLAAVRGEQW TPIEPKPRER PQIGGTIKQP PSNPPPRPPA EARKKPSEEE AAAAAGGPPG

[181]GPQVNPIPVT DEVV

Potential Transmembrane Residues

TRANSMEM [5]W to [28]T TRANSMEM [36]Y to [54]P TRANSMEM [90]A to [110]T

Know Mutations found in Autosomal Recessive CGD Patients

VARIANT [89]R -> Q (IN AR-CGD); VARIANT [93]H -> R (IN AR-CGD);

VARIANT [117]S -> R (IN AR-CGD); VARIANT [155]P -> Q (IN AR-CGD)

The shaded regions in the amino acid sequences listed for gp91-*phox* and p22-*phox* denote the potential transmembrane regions for the subunits deduced from hydropathy analysis. The residues involved in these proposed transmembrane domains are also listed. Proposed sites of glycosylation are designated by bold, double underlined, asparagine residues (**N**). All histidines in both gp91-*phox* and p22-*phox* are shown in bold (**H**). The residues implicated in the FAD binding domain, as well as the residues of the proposed NADPH binding sites are listed below the amino acid sequence for gp91-*phox*(90-92).