Overexpression and characterization of the cystic fibrosis protein CFTR using AFM and MALDI-TOF-MS
by David Guire Wooster

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biochemistry
Montana State University

Abstract:
A need exists for the development of additional methods capable of providing structural information about large transmembrane proteins such as the protein responsible for cystic fibrosis, the cystic fibrosis transmembrane conductance regulator, (CFTR), where traditional methods like NMR and X-ray crystallography have not been applicable. In the study reported here, CFTR was initially tagged with a poly(histidine) tail and overexpressed in a baculovirus system, allowing for the only one-step purification of CFTR that has been reported, using a nickel-affinity chromatography column.

Sufficient CFTR was obtained using this method to enable the mapping of 80 proteolytic peptides, identified by MALDI-TOF MS, onto the primary sequence of CFTR. This same method was then used to obtain 3D structural information concerning CFTR by exposing native CFTR, still within cellular membranes, to the hydrophilic, cysteine-alkylating reagent IAA and observing changes in mass. One cysteine located within the predicted pore region of CFTR, Cys-343, was found to be IAA-accessible and therefore accessible to the solvent. However, 3 additional cysteine residues, one each within the NBD1 domain, R-domain, and C-terminal tail region, were found to be inaccessible to IAA and therefore predicted to be buried within the 3D structure of CFTR. This is the first method concerning the use of mass spectrometry on whole CFTR in its native environment. In another study, AFM was used to characterize purified and liposome-reconstituted CFTR from this lab for the purpose of detecting protein-protein interactions involving CFTR. Here, the cytoskeleton protein F-actin was found to interact with CFTR without the need for accessory proteins. Additionally, this study constituted the first reported attempt at using atomic force microscopy to characterize protein-protein interactions involving CFTR, or any other transmembrane protein, in a lipid environment.

OVEREXPRESSION AND CHARACTERIZATION OF THE CYSTIC FIBROSIS PROTEIN CFTR USING AFM AND MALDI-TOF-MS

by

David Guire Wooster

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

of

Doctor of Philosophy

in

Biochemistry

MONTANA STATE UNIVERSITY-BOZEMAN
Bozeman, Montana

April 2002

APPROVAL

of a dissertation submitted by

David Guire Wooster

This dissertation has been read by each member of the dissertation 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.

Martin Teintze, Ph.D. _____ 4/19/02
(Signature) Date

Approved for the Department of Chemistry and Biochemistry

Paul A. Grieco, Ph.D. _____ 4-19-02
(Signature) Date

Approved for the College of Graduate Studies

Bruce R. McLeod, Ph.D. _____ 4-19-02
(Signature) Date

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University-Bozeman, I agree that the Library shall make it available to borrowers under the 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 in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Signature David G. Woods

Date 4/19/02

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

A need exists for the development of additional methods capable of providing structural information about large transmembrane proteins such as the protein responsible for cystic fibrosis, the cystic fibrosis transmembrane conductance regulator, (CFTR), where traditional methods like NMR and X-ray crystallography have not been applicable. In the study reported here, CFTR was initially tagged with a poly(histidine) tail and overexpressed in a baculovirus system, allowing for the only one-step purification of CFTR that has been reported, using a nickel-affinity chromatography column. Sufficient CFTR was obtained using this method to enable the mapping of 80 proteolytic peptides, identified by MALDI-TOF MS, onto the primary sequence of CFTR. This same method was then used to obtain 3D structural information concerning CFTR by exposing native CFTR, still within cellular membranes, to the hydrophilic, cysteine-alkylating reagent IAA and observing changes in mass. One cysteine located within the predicted pore region of CFTR, Cys-343, was found to be IAA-accessible and therefore accessible to the solvent. However, 3 additional cysteine residues, one each within the NBD1 domain, R-domain, and C-terminal tail region, were found to be inaccessible to IAA and therefore predicted to be buried within the 3D structure of CFTR. This is the first method concerning the use of mass spectrometry on whole CFTR in its native environment. In another study, AFM was used to characterize purified and liposome-reconstituted CFTR from this lab for the purpose of detecting protein-protein interactions involving CFTR. Here, the cytoskeleton protein F-actin was found to interact with CFTR without the need for accessory proteins. Additionally, this study constituted the first reported attempt at using atomic force microscopy to characterize protein-protein interactions involving CFTR, or any other transmembrane protein, in a lipid environment.

CHAPTER 1

INTRODUCTION

Cystic Fibrosis and CFTR

The most common lethal genetic disease in North America is cystic fibrosis, with approximately one in every 2,500 infants in the U.S being born with it (1). There are currently over 730 separate mutations within the cystic fibrosis gene identified as causing the disease (2), making carrier frequency for CF in the U.S. high enough that about one in every twenty Americans of Northern European descent are capable of passing on this autosomal recessive disease (1). The first known reference to this disease can be traced to folklore of the Middle Ages, where children's songs suggest midwives were able to notice a connection between salty sweat of infants in their care and their premature death. These afflicted children were often thought to be bewitched (3). Today, through the use of techniques such as microsattelite haplotyping, it has been found that the most common CF-causing mutation, ΔF508, arose over 50,000 years ago, and later became prevalent in the population perhaps as protection against many of the bacteria-induced diarrheal diseases such as cholera that have no doubt been frequent since the beginning of human civilization, due in large part to inadequate sanitation (4). The role of the cystic fibrosis protein in these bacterial-induced diseases is now known to be due to the gene product's importance in the regulation of fluid volume in the intestines, not unlike how this protein is involved in mucus hydration in the lungs. Although CF is a significant genetic

disease in North America, with just over 30,000 known cases in the United States alone, secretory diarrhea due to the overactive CF gene product expressed in the large intestine is the second leading cause of infant mortality in the developing world, estimated to be responsible for up to 3 million deaths per year of children under the age of 5 (5,6). Today, due in large part to the fields of molecular biology, physiology, and biophysics, the ancient disease marker for cystic fibrosis of salty-tasting sweat on a newborn infant's forehead is now understood to be caused by a defect in a chloride ion channel, the cystic fibrosis transmembrane conductance regulator, or simply CFTR.

The CF gene, which codes for CFTR, was discovered in 1989 by a lengthy process involving positional cloning (7). Fifty years before this discovery, in the 1940s, chronic bacterial infections in the lungs of CF patients were, for the first time, being recognized as a major contributor to CF mortality. In 1900, an infant born with CF could have been expected to live less than 5 years. Following the introduction of antibiotics and enzyme replacement therapy after WWII, life expectancy climbed significantly, but has since leveled off to where it is today, at approximately 30 years of age (8). A major disappointment in the field of CF research in the last 12 years has been the realization that, in spite of the knowledge of the exact location of the CF gene, no new effective treatments have emerged utilizing this hard-won information. This is unfortunate because cystic fibrosis is one of only a handful of diseases that can be attributed to mutation in a single gene and is therefore an ideal candidate for molecular therapies, including those involving protein and/or gene replacement. Despite often being described as one of the most studied proteins in the last ten years, structural information

concerning CFTR protein has been scarce, due in large part to difficulties inherent in the characterization of large transmembrane proteins like CFTR. Most significant among these obstacles are poor solubility, low and inconsistent yields using eukaryotic over-expression systems, and poor signal to noise ratios inherent in spectroscopic techniques (due to light scattering of lipid membranes). Additionally, transmembrane proteins are too large for NMR study, and are difficult to crystallize for X-ray diffraction studies.

Several theories have been put forward in an attempt to explain precisely how a defect in a single chloride channel could cause cystic fibrosis, with no single one of these theories being totally accepted. The most commonly repeated theory revolves around the notion that the CFTR chloride channel is necessary for the proper hydration of the mucus coating secretory epithelia, primarily those found in the lungs (9). Because tissues do not possess pumps for moving water from one compartment to another, plants and animals have instead come to rely on the formation of osmotic salt gradients to encourage the movement of water between various compartments. Normally, water in the airway surface fluid of the lung lumen finds its way there from the bloodsteam by following sodium chloride gradients set up by ion channels as shown in figure 1. One of the most vital of these ion channels is CFTR. When CFTR is not found at the apical membrane of lung epithelial cells in active form, such as what occurs in CF, water cannot move from the bloodstream and into the airway surface fluid of the lung to hydrate the mucus and allow it to be removed, along with dust and microorganisms, in a timely fashion. Instead, the mucus becomes overly viscous and difficult to remove from the

lungs, setting the stage for the well-known chronic bacterial infections and inflammation that afflict CF patients during most of their short lifetimes.

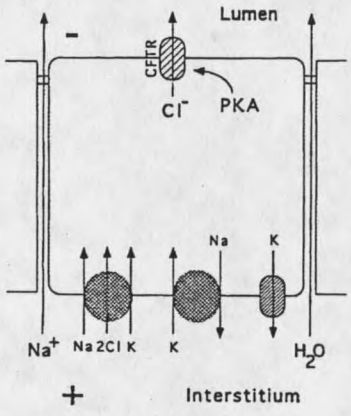

Figure 1: **Secretory epithelial cell** (e.g. lung, intestine) indicating the role of CFTR in directing net H_2O flow from the interstitium (bloodstream) into the lumen for hydration of the mucus layer (not shown). Note that Na+ flows around cells in order to neutralize negative charge built up by Cl- flow through cell via the CFTR chloride channel. **PKA**: a kinase activated by increases in cAMP. When activated, PKA phosphorylates serines within the R-domain of CFTR triggering channel opening. The driving force for Cl- secretion is the build-up of negative charge due to the flow of K+ out of the cell via the potassium channel on the basolateral side of the cell.

Search for the CF Gene

In the late 1980s, two collaborating groups, one at the University of Toronto and the other at the University of Michigan, were closing in on the location of the CF gene on the long arm of human chromosome 7 using RFLP markers obtained from families known to be carriers of CF. These researchers were aware that the technique they were using, chromosomal walking, carried with it the inherent limitation that it was impossible to narrow down the putative CF gene to less than 5 total "candidate genes" on chromosome 7 in the area between their closest RFLP markers (7). A hypothesis was therefore needed to help determine which of these 5 genes on human chromosome 7 was the CF gene. It was therefore reasoned that any single gene capable of causing a disease as devastating as CF by itself was likely to be conserved among various mammals. One way of testing this hypothesis meant the generation of a series of Zoo Blots (or, Southern Blots performed on genomes across several species) of the genomic DNA of a diverse group of mammalian species (10). Restriction digests of these mammalian genomic libraries were probed with short nucleotide sequences based on each of the 5 candidate human genes. It was soon clear from the Zoo Blotting analysis that only one of the 5 genes was located within each of the mammalian genomes tested. This gene turned out to be a 250 KB stretch of DNA on the long arm of chromosome 7, band q31, and became the new focus of their work (7). The next step towards proving that this was indeed the CF gene entailed obtaining full-length cDNA copies of the mRNA from a subject without CF and then performing Northern Blotting analysis for the purpose of locating the

expression pattern of the gene in various human tissue samples. It was reasoned that if this unknown gene was the CF gene, then mRNA corresponding to it should be detected in tissues affected in cystic fibrosis, such as the epithelial tissue of the lungs. Once this was verified by Trezise et al. in 1991 (11), full-length cDNA representing the mRNA from the proposed CF gene was tested for its ability to change the properties of an immortalized airway epithelial tissue culture (from a CF patient) back to the wild-type state. This was shown to be the case using retroviral vectors, accompined by patch clamping procedures, providing further confirmation that CFTR was indeed the long sought after CF gene (12,13). Within two years, peptides would be synthesized based on the extracellular loops of CFTR, and antibodies specific to these loops generated for the purpose of detecting the subcellular localization of the CFTR protein, which turned out to be localized to the apical, but not the basolateral, membranes of secretory epithelia in the lungs, where it was expected (14,15).

CFTR: Structure and Function

CFTR is a 1480 amino acid transmembrane glycoprotein, widely expressed in several secretory and absorptive epithelial tissues, including those of the pancreas, lung, large intestine, sweat duct, and vas deferns (16). CFTR is one of only a few members of a widely distributed superfamily of proteins called the ABC Transporters (for **A**TP **B**inding **C**assette) that is not itself a transporter, but rather an ion channel. CFTR also carries with it the distinction of being the only known ABC superfamily member that has 5 separate domains as opposed to 4 domains characteristic of all other ABC proteins, the

5th domain being the "R-domain" (so-named because its function appears to be in regulation of channel gating, along with two NBD domains)(16). However, CFTR is similar to other ABC family members in that it possesses two nucleotide binding domains (NBD1 and NBD2), both of which bind and hydrolyze ATP. CFTR also has two transmembrane domains, each consisting of 6 predicted transmembrane helices (17).

The CFTR gene itself appears to be a relatively ancient gene, believed to have arisen as the result of a gene duplication event of an unidentified ancestral prokaryotic ABC transporter gene prior to the divergence of fishes (18). This gene duplication event was, in all likelihood, then followed by fusion of the duplicated halves into a tandem arrangement along chromosome 7 forming a mirror image of the original transporter gene within the entire, newly fused-together CFTR gene (Figure 2). Some evidence for this

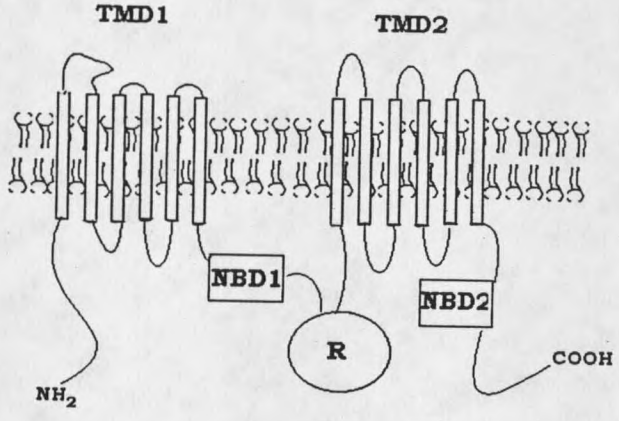

Figure 2: **Predicted topological structure of CFTR** showing the 5 domains. The cytoplasmic portion consists of the R-domain and the two nucleotide binding domains (NBD1 and NBD2), all 3 of which have been shown to be involved in channel gating. Each transmembrane domain (TMD) consists of 6 predicted helices. The most common CF-causing mutation is a deleted Phe-508 in NBD1.

Some evidence for this duplication event followed by a fusion event includes the finding of significant sequence similarity between the two halves of CFTR, as well as the fact that all other ABC transporter proteins not believed to have arisen as the result of gene fusion appear to function normally as dimers with one another (18).

While nearly all ABC transporters appear to use both NBD domains to hydrolyze ATP and transport a wide variety of substances including amino acids, peptides, proteins, small hydrophobic molecules (e.g. the multi-drug resistance protein Pgp), lipids such as cholesterol and phospholipids, sugars, inorganic ions, and polysaccharides (18) against a concentration gradient, CFTR on the other hand appears to have, at some point in its history, managed to "commandeer" its NBD domains into functioning during the process of "gating" of its anion channel, apparently as an added layer of regulation during opening and closing of the channel. One line of evidence for this is that the measured V_{max} of ATP hydrolysis by the CFTR NBD domains matches closely the measured kinetics of CFTR channel opening and closing as determined by patch clamping (~1/sec) (19).

Site-directed mutagenesis has been extensively performed on the CFTR sequence during the last decade, and much of the evidence accumulated from this has suggested that amino acid changes localized to both the NBD and R domains tend to alter the kinetics of channel gating (20, 17), while amino acid changes in the predicted transmembrane domains affect anion channel selectivity as well as conductance, the most likely explanation for this being that the transmembrane domains are where the pore is located (21, 22, 23). Three of the loops of CFTR which, like the other 7 loops, function

to join the transmembrane helices with one another, have also been selectively mutated to produce ion channels having alterations in both halide ion selectivity as well as changes in channel gating kinetics (24). Exon 13 of CFTR encodes the R domain. This exon was inserted into the ancestral CFTR gene following fusion of the duplicated halves as discussed above. Curiously, the R domain of CFTR, while having no known sequence homology to any other eukaryotic or prokaryotic protein domain, has been found by Dulhanty et al. to possess sequence similarity to DNA polymerases of viruses (25). The significance of this finding is still unknown. The CFTR gene also possesses the remnants of L1 retrotransposons within the intron sequences flanking exon 9. Exon 9 encodes a highly conserved section of the NBD1 domain of CFTR, and at various times during human evolution, these L1 elements near exon 9 of CFTR are theorized to have undergone transcription from the opposite direction of the gene. These transcripts were then reverse transcribed into a cDNA copy, and finally co-integrated along with the L1 sequences where they eventually came to rest within at least 10 different locations throughout the human genome, effectively "amplifying" CFTR exon 9 sequences within the human genome (26). This basic mechanism of exon shuffling via the reverse transcription of L1 elements is believed to have occurred in the evolution of many other human genes as well. Also of interest, the two NBD domains of CFTR have been by found by Pratt et al. to contain sequence similarity to a known actin-binding protein (see chapter 4) (27).

The NBD domains of ABC transporters are the most conserved portions of the protein (30-50% similarity), each consisting of a short Walker A and Walker B motif specialized

for the binding and hydrolysis of ATP, respectively (28, 29). Perhaps as an indication of their importance for proper ion channel function, most CF-causing mutations occur in these two NBD domains (30). In addition, there appears to be significant similarity between the NBD domains of CFTR and G-proteins, although threading of the primary sequence of the NBD1 domain of CFTR onto the crystal structure of the G-protein Ras was not considered particularly revealing. However, the glycine in the LSGGQ "signature sequence" (so-named because this sequence, or one like it, is found in all known ABC superfamily members) within the NBD domains of CFTR corresponds to a conserved glycine in G-proteins and is necessary for GTP hydrolysis (31). The presence of this short LSGGQ signature sequence near the Walker motifs of the NBD domains also serves to distinguish the ABC superfamily of proteins from other known ATP-binding protein families (e.g. G-proteins, ATPases, and kinases), and this signature sequence has been shown to be important for the transduction of free energy resulting from ATP binding and hydrolysis (which takes place at the Walker motifs), into mechanical energy for the pumping of substrates (or ion transport gating in the case of CFTR). To date, there are 4 crystal structures of NBD subunits of ABC family members in the literature (Figure 3), and these structures point to a conserved mechanism for the binding and hydrolysis of ATP in ABC transporter family members, as well as the transduction of this energy into opening and closing of the pore region [MalK (32), Rad50 (33), HisP (34), and MJ0795 from Methanococcus jannaschii (35)]. At near-atomic resolution, crystallized NBD domains can be seen as forming an overall "L" shape. The NBD domains have both Walker A and B motifs (and therefore the site of

ATP binding and hydrolysis) contained within one arm of this "L" structure and a conserved LSGGQ sequence within the other arm. The "LSGGQ arm" contained within the "L" structure has been theorized to interact, via a conformational change in the overall structure of the NBD domain during ATP hydrolysis, with the putative pore region formed by the transmembrane helices of CFTR (34). (Figures 3 and 4)

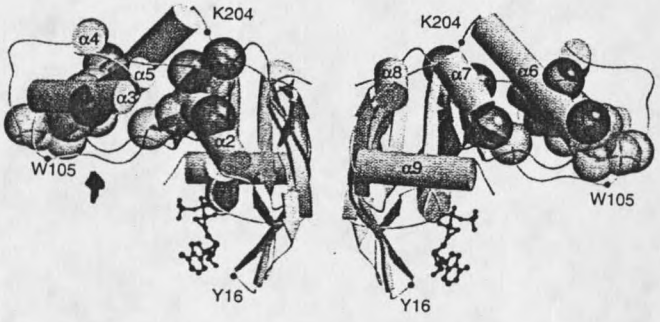

Figure 3: **Crystal structure of HisP NBD dimer** (34). HisP NBD domain was the first ABC superfamily member NBD to be crystallized. Arrow indicates relative position of Phe-508 in CFTR NBD1. Each monomer has an "L" shape that is believed to change shape during ATP hydrolysis, allowing it to interact with the pore region of the transporter during pumping.

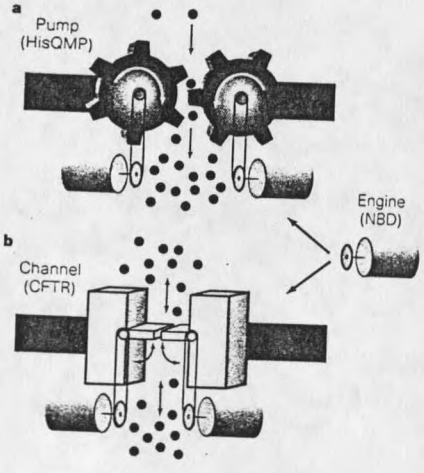

Figure 4: Two conceptual diagrams describing the role of NBD domains in transporters such as HisP (top) verses channels like CFTR (bottom). NBD domains appear to function mostly as "gates" in ion channels like CFTR while in transporters (HisP) they may function more as pumps (34).

Recently, the total structure of MsbA, a bacterial ABC transporter was solved using X-ray crystallography, as well as the structure for P-glycoprotein to 10 Å using 2-D crystals and cryoEM (182, 188).

CFTR can also be distinguished from other ABC Transporter family members, such as the multi-drug resistance protein P_{gp}, in being the only ABC protein exhibiting "functional asymmetry", a situation whereby one half of the protein is incapable of replacing the other half of the protein without causing changes in overall ion channel function (28). In part because of this finding, it is now theorized that each of the two NBD domains of CFTR is involved in separate functions during channel gating. NBD1 is the most extensively described NBD of the two and has been shown to be involved with channel opening during the gating process, while the second NBD domain, NBD2, is more closely associated with ion channel closing (36, 37). Non-hydrolyzable ATP analogs along with mutagenesis were also involved in studies used to bolster these findings (28, 38, 39).

The primary amino acid sequences of the transmembrane domains of ABC proteins has not, to date, allowed the prediction of what specific substrates a given ABC protein will transport through its pore. Naturally occurring CF mutations within CFTR, as well as site-directed mutagenesis within the transmembrane domains during the last 10 years, have led to the conclusion that 6 positive charges localized within the area of the membrane (2 highly conserved lysines and 4 arginines in these predicted helices) influences, to varying degrees, the selectivity sequence of anion travel thru the pore, as well as the rate of ion travel (i.e. conductance) (7).

Probably the first paradigm proposed to explain the overall mechanism of channel gating of CFTR was based on the previously well-studied gating kinetics of Shaker, a *Drosophila* potassium ion channel found in neurons (40). Using this analogy, the R-domain of CFTR can be compared to the amino terminal ball domain of Shaker. Shaker is known to shut off passage of potassium through its pore via its N-terminal ball domain, which is free to move into the pore region from the cytoplasm. This "plugging" of the

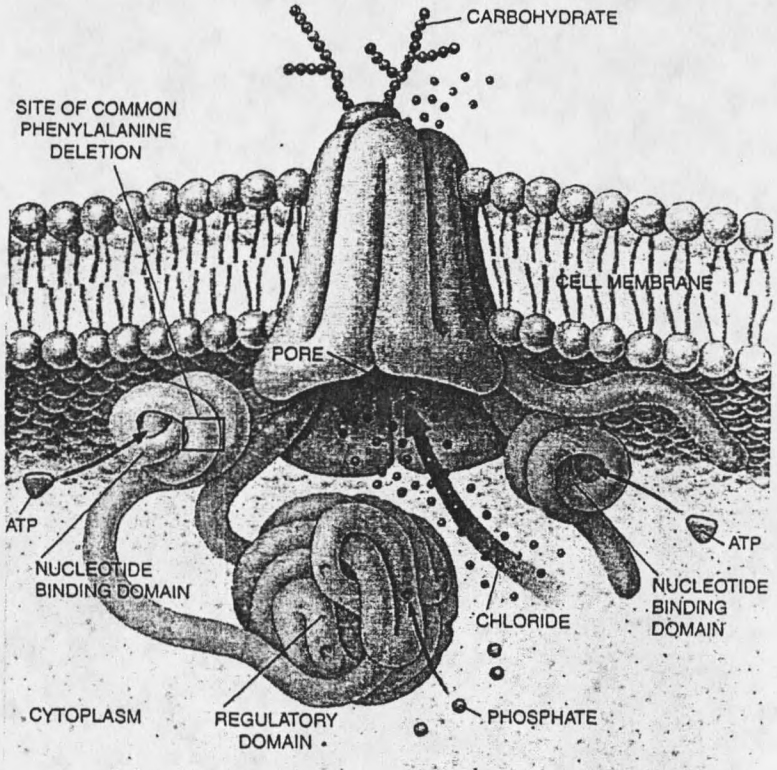

Figure 5: **Cartoon drawing of CFTR**. Channel is gated open for chloride flow out of cell (124).

pore by the ball domain of Shaker has been found to occlude ion travel after repolarization of the neuronal membrane following the action potential (40). The R-domain of CFTR, on the other hand, is normally post-translationally modified (phosphorylated by PKA at 9 serines on its R domain) during channel opening (41). When the R domain is dephorphorylated (by constitutively active phosphatase enzymes co-localized in the membrane), channel closing occurs in CFTR, perhaps by a similar mechanism of closing as the Shaker channel by its amino terminal ball (40). Mutagenesis experiments in which much of the R-domain of CFTR has been deleted have served to verify only a portion of this paradigm, however (17).

More recently, a second paradigm has been suggested based on the notion that the R-domain of CFTR functions in a way similar to recently discovered G-protein exchange factors (GEFs), while the NBD domains of CFTR would be analogous to a pair of G-proteins (28). In this analogy, the R-domain enables the NBD domains of CFTR to pick up ATP from the cytoplasm in much the same way that a G-protein picks up GTP more efficiently when a GEF is present to interact with it. And the pore region (probably located in the transmembrane domains of CFTR) would then be analogous to the G-protein's downstream effector (usually an enzyme), causing the pore, located within the transmembrane domains, to open when the now ATP-bound NBD activates the pore region (analogous to when a G-protein is bound to GTP and activates an effector enzyme downstream). Also, the NBD2 domain, or perhaps a dephosphorylated R-domain, both of which are known to be involved in channel closing, could then function as a GAP (GTPase-activating protein), causing the NBD1 domain to hydrolyze ATP and release

ADP and P_i, thereby shutting down the channel. In summary, this intriguing analogy carries with it the potential for describing CFTR as a kind of "miniature signal transduction system", self-contained within the same 168 kd transmembrane protein.

In the first study reported here (chapter 3), the nucleotide cDNA of CFTR was modified by the addition of 10 histidine codons and a stop codon onto the C-terminal end. This was accomplished in part with the use of PCR. The resulting modified CFTR cDNA was then used to generate a recombinant baculovirus, capable of infecting insect cells and encouraging them to produce the CFTR protein in relatively large amounts. This procedure allowed for a new method of purification of CFTR using an affinity-based chromatography rather than the more subjective and difficult to reproduce hydroxyapatite chromatography.

In the second study (chapter 4), CFTR was expressed, purified, and reconstituted in to a liposome and sent to Boston for analysis using atomic force microscopy at Boston University and Harvard Medical School. Here, the chloride channel activity of CFTR was measured using a highly sensitive technique called "patch clamp". In addition, this CFTR was subjected to the surface-imaging technique AFM for the purposes of determining possible interactions with the cytoskeletal protein actin.

Finally, in the last study (chapter 5) purified whole CFTR was proteolyzed for the purposes of identifying the resulting peptides using the mass spectrometry technique MALDI-TOF MS, and the detection of any covalent modifications made by chemical reagents. With this procedure, the weights of individual peptides can be established often to a resolution of 0.01 Daltons, allowing for unambiguous identification of the

peptides being tested. In addition, it is possible to detect the accessibility to the solvent of various residues of a protein by the detection of a mass change. This was also accomplished for one Cys residue, Cys-343, which is located within an area of the protein long suspected to line the pore of the CFTR ion channel.

CHAPTER 2

PREVIOUS ATTEMPTS AT CFTR PROTEIN EXPRESSION AND CHARACTERIZATION

Introduction

Obtaining mammalian transmembrane proteins in amounts sufficient for structural characterization is difficult. Therefore, it is often necessary, as a first step in this process, to overexpress the gene that codes for the foreign protein in a cell line which is easier to grow and more efficient at producing the protein than the endogenous source. Many important functional studies would otherwise be impossible without heterologus expression systems using cell lines such as bacteria, insect, yeast, amphibian, or mammalian cells. In the decade since the discovery of CFTR, the first chloride channel cloned, as well as the first human disease-causing gene discovered by the process of positional cloning, numerous attempts at expression and characterization of CFTR have been made. Reviewing some of these important studies provides for a context in which to consider the studies discussed now.

Literature Review

Once the gene for CFTR became available in 1989, investigators from a surprisingly diverse number of fields managed to find a way to incorporate CFTR into their work. As is often the case with protein expression studies, attempts at the overexpression of CFTR began (within a year of its discovery) with *E. coli* in 1990, when Gregory et al.,

produced the CFTR protein in this bacterium (42). Disappointingly, human CFTR produced by *E. coli* was found to be improperly folded, and was not obtained in large enough yields. Purificaton was a problem, and no functional assay was reported using this bacterially expressed CFTR. However, proteolysis fingerprinting was performed on CFTR by this same group, showing that the recombinant CFTR protein made in *E. coli* was indistinguishable from native sources of CFTR, encouraging future heterologous CFTR protein expression in other systems. This was also the first attempt at using polyclonal as well as monoclonal antibodies to precipitate and purify CFTR. Also in 1990, Gregory et al. reported using an in vitro translation system to express CFTR employing reticulocyte lysate in the presence of S-35 methionine. Using autoradiography following gel electrophoresis, it was shown that the CFTR protein obtained by this in vitro translation system migrated with the apparent molecular weight of 135 KD in an SDS-PAGE gel, where CFTR's closest homologue, Pgp, was known to migrate (42).

In 1991, further attempts were made to purify CFTR, one involving the use of an affinity column bound with triazine dye as a mimetic for ATP. CFTR protein was found to bind too tightly to the column via its NBD domains to be eluted (43). Ostedgaard and Welsh later used an affinity column consisting of a bound lectin, specifically, wheat germ agglutinin. In this case, CFTR was shown to bind avidly to the lectin via its N-glycosylation sequences, but unlike with the triazine dye column, CFTR was this time successfully eluted off using excess soluble carbohydrate in the buffer.

Using this purification system, CFTR was reported to be enriched 8-9 fold over native membrane proteins after the initial purification step.

Over-expression trials reported in 1991 included the use of a diverse number of cell types, including amphibian, as well as mammalian-based cell systems (44). Fortunately for many ion channel activity investigators, obtaining large amounts of purified CFTR is often not a necessary step when using amphibian cell lines like Xenopus oocyte cells, as CFTR channel activity is readily assessed in Xenopus oocytes without the need for any time-consuming and costly isolation procedures. This is accomplished by using extremely sensitive electronic patch clamping methods after introducing CFTR cDNA or mRNA into the cells. Patch clamping has an advantage in that it can be performed on cell membranes without the need for purification of the CFTR ion channels from the cell membranes, and is currently one of only two techniques with resolution high enough for the measurement of the activity of a single protein molecule (the other being single molecule fluorescence spectroscopy). At about this same time, studies using CFTR cDNA containing naturally occurring CF-causing mutations, such as in the NBD domains, helped to confirm what had been previously theorized: that the severity of the mutation affecting chloride conductance in CFTR parallels the severity of disease that the patient presents with, providing further evidence that mutations causing the lowest chloride conductance are correlated closely with the most severe cases of CF (45). Another significant finding from studies using naturally occurring CFTR mutants was that some of them were found to be activated by stimulators of intracellular cAMP such as forskolin, not unlike the way CFTR is normally activated in tissues CFTR is

endogenously expressed in, providing evidence that in heterologus systems, CFTR can be expected to function similarly to wild-type CFTR, and activated in a similar fashion to CFTR in its native epithelial tissue (45).

Expression of CFTR in mammalian cell-based expression systems would prove useful in "coexpression experiments", where it was found that CFTR has the surprising capability of directly regulating other ion channels, including the ion channel ORCC (outwardly rectifying chloride channel). Typically, in these studies the mRNA or cDNA coding for CFTR is injected into IB3-1 cells (a CF bronchial epithelial cell line) known to expresses ORCC. Then, using patch clamping techniques, it can be shown that ORCC is upregulated by the NBD1 domain of CFTR, lending experimental support to the predicted "regulator" function indicated in the name cystic fibrosis transmembrane conductance regulator (46).

The development of an animal model mimicking the human form of a disease is almost always a necessary step on the road towards the understanding the underlying etiology of a disease. Two years following the discovery of the CFTR gene in 1989, a portion of CFTR exon 10 was mutated into a stop codon, making the CFTR protein nonfunctional in embryonic stem cells of mice by Koller et al., an important step towards generation of a cystic fibrosis CFTR knockout mouse (47). One year following this work, generation of a CFTR knockout mouse was reported by Snouwaert et al. Importantly, CF mice homozygous for the disrupted CFTR gene were shown to display many of the features common to cystic fibrosis patients; unfortunately, these transgenic mice died prematurely due to intestinal obstructions (48). This problem has since been

corrected by the addition of laxative supplements in the food of the mice and, as a result of this, as of the mid-1990s most CFTR knockout mice have a lifespan long enough to develop many of the same types of the lung infections and inflammation characteristic of the cystic fibrosis lung (49).

The CFTR gene has been incorporated into the genomes of several viruses, often for the purpose of gene therapy. Some of these viral vectors include: adenovirus, adeno-associated virus (AAV), baculovirus, vaccinia virus and retrovirus. The goal of gene therapy for CF is the restoration of proper mucus hydration resulting from the replacement of the CFTR protein in secretory lung epithelia. To date, no significant improvement in the condition of any CF patient has been reported with gene therapy involving the CFTR gene (phase I and phase II trials only), probably due to low expression efficiency, however several studies are still ongoing (50).

Expression of a foreign gene in a heterologus cell line can be achieved under transient as well as stable conditions, depending in part on how the selection of transfected cells is accomplished. In 1991, CFTR cDNA was stably integrated into a cell line not normally expressing CFTR by Rommens et al. (51). This group obtained production of the CFTR protein, both wild-type CFTR as well as ΔF508 mutant CFTR, in a stably integrated mouse fibroblast cell line. Verification of CFTR expression in these studies was obtained by measuring a concomitant increase in the characteristic cAMP-induced chloride conductance when the cells were transduced with wild-type, but not ΔF508 CFTR cDNA, as would be expected if the ΔF508 CFTR was either too rapidly degraded or inadequately transported to the cell membrane.

In 1992, an important proof-of-concept experiment involving gene therapy was reported when the cDNA of CFTR was transfected into cultures of immortalized epithelial cells from a CF patient. Surprisingly, it was found that as little as 6-10% transfection efficiency of this human epithelial cell line cultured in such a way as to grow in a similar manner to a lung epithelia (forming gap junctions, etc.) was needed in order to observe a return to wild-type properties of this tissue culture system, providing encouragement for gene therapy treatments (52).

What may be considered the first overexpression of CFTR protein, in a practical sense, for the purposes of purification and reconstitution, was achieved by Bear, et al. using the baculovirus expression system in 1992 (53). In the last 15 years, baculovirus expression systems have been used for the production of a diverse number of mammalian proteins from many sources, and just as importantly, in relatively large amounts. CFTR protein expression of 0.5 mg/L insect cell culture was obtained. Bear, et al. also developed, in 1992, a purification procedure for purifying CFTR from baculovirus-infected Sf9 cells using hydroxyapatite chromatography. Following a gel exclusion chromatography step, purified CFTR from was then reconstituted into proteoliposomes and tested for chloride channel activity with an electronic method similar to patch clamping, except adapted for CFTR incorporated into planer lipid bilayers. This study established unambiguously for the first time that pure CFTR by itself, without the possible assistance of other proteins, functions as a chloride channel (53). Ramjeesingh et al. later engineered a poly-histidine tag into this CFTR construct to aid in purification (54).

In 1992, Fuller et al., using polyclonal antibodies, precipitated "CFTR-related polypeptides" from mammalian cell lines, including HeLa, Bsc-40, HEp-2, and Chinese hamster ovary cells. These cell lines were transfected with CFTR cDNA using the vaccinia T7 protein expression system. Fuller et al. also reported obtaining similar results using a yeast-based expression system (55). By the mid 1990s, CFTR was being produced in both mammalian 293 and COS-1 cells. Seibert et al. would later go on to express the CFTR cDNA as a transient construct under the control of CMV promoter. Calcium phosphate was used for transfecting these mammalian cells, and pure CFTR was subsequently harvested approximately 48 hours post-infection (56).

Since the late 1980s, transgenic plants and animals have become an option for the over-production of foreign proteins in addition to more traditional, *in vitro* cell lines. CFTR was successfully over-expressed but not reportedly purified using the milk of transgenic mice by the biotechnology company Genzyme in 1992. CFTR was found in this study to be associated with the fat globules in the milk of these mice, and CFTR was said to be present in high enough yields as to be of use in protein replacement therapies, with the added provision that these results could be repeated in larger animal trials (e.g. goats). Unfortunately, nothing further about this or any other transgenic animal or plant over-expression method for CFTR (57) since 1992 has been reported, no doubt having the overall effect of discouraging future attempts at protein replacement therapies for CF.

When attempting structural studies on large transmembrane proteins having several domains, it is often necessary to adopt methods bypassing the twin problems of low yield and inherent lack of solubility. In 1992, the NBD1 domain of CFTR was produced as a

separate, soluble polypeptide in *E. coli* bacteria by Ko et al (20). The NBD1 domain, which has been implicated in playing a central part in ion channel gating, was produced in functional form as a soluble fusion protein, in frame with maltose-binding protein (MBP) both as the wild-type construct as well as the deltaF508 mutant version of NBD1 (58). These investigators engineered the CFTR-MBP chimera construct to, in their own words "aid in purification, solubilization, and crystallization", with the final goal of studying the effects of protein-protein interactions between this domain and others of CFTR. Both the wild-type NBD1 fusion construct as well as the ΔF508 NBD1 were found to be functional by use of a binding assay involving a fluorescent nucleotide analog of ATP. The recombinant NBD1 polypeptide's alpha-helical and beta-sheet secondary structure was then assessed using UV- circular dichroism. That same year, Hartman et al. also reported expressing NBD1 as a soluble domain using, as a criteria for functionality, ATP-agarose binding assay, as opposed to the fluorescent nucleotide analogs used by Ko to ascertain ATP binding activity (59). Similarly, 3 years later, the second NBD domain, NBD2, was expressed and purified as a separate soluble protein by Randak et al. (also from *E. coli*), whereupon the K_d of NBD1 for ATP binding was reported to be accurately measured for the first time (μM range) (60).

Teem et al. in 1993 constructed a yeast-human chimeric membrane protein consisting of the NBD domains of human CFTR in place of the endogenous NBD domains of the yeast mating factor Ste6 (61). Despite replacement of its endogenous NBD domains with the NBD domains from the ion channel CFTR, the yeast Ste6 transporter was still able to export mating factor normally against a concentration gradient. However, mating

efficiency was significantly reduced when the ΔF508 version of NBD1 was used instead of wild-type NBD1.

Refinement of insect expression systems using baculoviruses specialized for the expression of CFTR appears to be an ongoing process. In one case involving affinity chromatography in 1993, Peng et al. was able to produce a full-length CFTR protein fused to glutathione-S-transferase in insect cells. The fused enzyme was later cleaved away from recombinant CFTR following purification (62).

In 1994, Chang et al., engineered exogenous glycosylation sites throughout the sequence of CFTR for the purpose of verifying topologically, all of the positions of each of its predicted loops as being either extracellular or intracellular. By assessing the glycosylation status of each of the predicted loops of CFTR, it was determined that the positions of each of the loops of CFTR was as originally predicted in the topological structure proposed in 1989, based solely on the primary sequence of CFTR (63).

As has often been the case with ion channel studies, isolation and purification is not always a necessary step on the road towards characterization of their structure and function. Along these lines, Tilly et al. measured CFTR chloride channel activity with electrophysiological methods (patch clamping) following overexpression of CFTR in the cell membranes of CHO cells and NIH 3T3 fibroblasts simply by isolating the cell membranes following expression, and fusing these "membrane vesicles" to artificial planar lipid bilayers (64). Mouse mammary carcinoma cells were then used by Reisin et al. in 1994 to stably express CFTR, whereupon electrophysiological studies performed

on membrane patches and whole cells revealed that CFTR transports ATP molecules as well as chloride (65), although these findings have been disputed (66).

In 1994, Sheppard et al. produced truncated CFTR protein (D836X) in Xenopus oocytes. This construct lacked almost the entire C-terminal half of CFTR and yet was still able to form an ion channel. This ion conductance was attributed to the formation of functional dimers, as alluded to in the introduction (67). The measured channel activity was slightly lower than the wild type CFTR, however. Given the overall symmetry in the topology between the first and second halves of CFTR, as well as the fact that many ABC transporter family members (and all ABC proteins from prokaryotes and eukaryotic intracellular organelles) have been shown to be function as homodimers, this is probably not surprising (68).

Anytime a foreign protein is produced in a new cell line, the lingering question of uniformity of function needs to be addressed. In 1995, O'Riordan et al. compared CFTR made in sf9 insect cells (via the baculovirus system) with CFTR expressed in CHO cells and found that, under the control of metallothionein promoter, CHO cells yielded only ~ 0.1% of their total cellular membrane protein as CFTR, however CHO and Sf9 CFTR performed nearly identically in electrophysiological properties when the two preparations were compared (69). O'Riordan et al. also reported purifying CFTR expressed in CHO cells using the detergent lysophosptatidyl choline (LPC) and DEAE ion exchange chromatography. LPC detergent was not capable of solubilizing CFTR from insect cell membranes, perhaps due to the incomplete glycosylation of CFTR resulting in a less soluble insect-produced CFTR. It has also been suggested by some

that CFTR may be associated with the cytoskeleton in insect cells, as this would also help explain the difficulty in solubilization (70).

It has been estimated previously that as many as 90% of all CF patients carry at least one allele of a temperature-sensitive folding mutation in CFTR called ΔF508. Therefore, experiments during the last 10 years have been attempted in an effort to characterize the intracellular movements of this particular deletion mutant with the eventual goal of somehow making more CFTR available at the cell surface of CF patients. Using mammalian expression systems, it was found that the ΔF508 version of CFTR becomes "stuck" in the ER, and as a result associated with the ubiquination degredation pathway, and eventually degraded. In 1995, pulse-chase experiments involving recombinant CFTR in CHO cells revealed that the ΔF508-CFTR protein tracks intracellularly the same way in CHO cells as it does in native human epithelial cells where CFTR is normally found. Because of this finding, expression of CFTR in CHO cells has provided investigators with a valuable tool for the characterization and involvement of chaperones and proteasomal degradation pathways used by mammalian cells in CFTR trafficking (71).

BHK (baby hamster kidney) cells with a stable integration of the CFTR gene provide a more uniform source for CFTR than CHO cells, however BHK cells are not as convenient for multiple applications, unlike the more versitile CHO cells. CFTR has been modified with a polyhistidine construct coding for 12 histidine residues onto the C-terminal end and expressed in BHK cells (72), Sf9 cells (73), and yeast (74).

Molecular biology has provided many useful techniques to protein chemists over the last two decades. One of the most useful of these from the standpoint of the characterization of CFTR has been the ability to modify proteins with site-directed mutagenesis. In 1995, Howard et al. engineered epitopes endogenous to the influenza virus (specifically, HA and M2 antigenic sequences) into the 4th extracellular loop of CFTR (75). These "virally-tagged" CFTR constructs were then produced in HeLa cells and the resulting CFTR tracked intracellularly using monoclonal antibodies to the viral epitopes. CFTR was tagged with these epitopic sequences at the C-terminus as well as in the fourth external loop. Only one of the constructs, the M2-modified CFTR, made it to the surface and found to retain full activity. In 1995, the Genzyme Corporation included the entire influenza transmembrane protein, HA, along with CFTR in a proteoliposome to encourage fusion of CFTR-containing liposomes to human epithelial cells for use in possible protein replacement therapy. The CFTR in these experiments had been co-reconstituted along with HA into artificial membranes, forming the potentially therapeutic proteoliposome with the goal of protein replacement therapy in the lungs of CF patients (76).

In work similar to the soluble polypeptide NBD1 domain studies in 1992, the R-domain of CFTR was expressed as a separate, soluble polypeptide in *E. coli* in 1996 by Townsend et al., for the purpose of identifying specific amino acid side chains on the R-domain normally phosphorylated by cellular protein kinase A (PKA). Following phosphorylation by endogenous PKA, purified R-domains were then subjected to proteolysis with trypsin, and the resulting peptides were subjected to mass analysis using

HPLC-ES-MS in an effort to ascertain sites of phosphorylation. Specific peptide masses differing in molecular weight by one phosphate group addition on the R-group peptides were found by comparing to theoretical weight of the unphosphorylated peptides. It was determined that only 9 out of a possible 14 predicted serines/threonine phosphorylation sites were phosphorylated *in vivo* on the R-domain following cAMP stimulation (77). A second study involving a soluble R-domain polypeptide produced in E. coli include one in which the R domain was injected exogenously into previously-activated CFTR expressed in Xenopus oocytes. These R domains were then shown to inhibit chloride channel current during patch clamping experiments, indicating that the R-domain was able to block the flow of chloride out of the CFTR pore when the R-domain was unphosphorylated (78). In another experiment, serine amino acids located in the R-domain in intact CFTR were individually mutated into negatively charged aspartic acid residues. Chloride channel activity was then assessed using patch clamping. These results demonstrated that the R-domain of CFTR appears to behave in the same manner regardless of whether the negative charge is the result of phosphorylation of serines by PKA or of mutagenesis of serines into acidic residues, suggesting that the overall charge of the R-domain is what determines whether it is able to deactivate the CFTR ion channel (79). It should also be noted that experiments have been performed on purified R domains which suggest that a conformational change takes place in the R domain due to phosphorylation, suggesting this modification functions to change the overall shape of the R-domain, perhaps causing it to bind to the NBDs. Both Pichiotto and Dulhanty observed mobility changes during electrophoresis, as well as changes in CD spectrum of

the soluble R-domain, and have therefore attributed them to changes in R domain conformation (80, 81). It is possible that the R-domain gates the CFTR channel as a result of changes in both overall charge as well as changes in conformation, as neither theory is mutually exclusive of the other. (Figure 4)

The baculovirus expression system, while it has proven to be the most efficient system in terms of yield of active CFTR, can be expensive and non-reproducible at times, depending on the protein. Difficulties associated with the baculovirus expression system account the increased popularity of the yeast *Saccharomyces cerevisiae* for the production of mammalian proteins. Single-celled fungi such as this can be grown in large volumes and to relatively large densities with less labor and with less expense than Sf9 insect cells require using the baculovirus system. In 1996, a group at the University of North Carolina-Chappel Hill used *Saccharomyces cerevisiae* to express functional CFTR, however the overall yield of CFTR was just 5-10% of that in Sf9 insect cells (82). They did note that the ease of use and low cost of yeast compared to insect cell lines made up for lack of yeast CFTR yield on a per-cell basis. Later, this group included a poly-histidine tag on the C-terminal end of their yeast-produced CFTR in 1998 to aid in purification (74).

In 1997 Chen et al. employed yeast artificial chromosomes (YACs) to introduce the entire 230 kb gene, including introns as well as flanking sequences, of CFTR into mammalian cells via liposomes (83). Also in 1997, Mogayzel et al. used a YAC with the entire human CFTR gene as well as adjacent sequences to the gene to produce CFTR in CHO cells, which helped facilitate gene regulation studies. The YACs in this case

had become integrated into the CHO cell genome with the resulting single integration event verified using FISH (84).

In 1998, Johnston, Ward, and Capito transfected CHO and HEK (human embryonic kidney) cell lines with the goal of over-expressing CFTR in amounts high enough to form large molecular weight "aggresomes". They described these aggresomes as "ubiquitin-rich cytoplasmic inclusions containing detergent insoluble, misfolded protein aggregates of CFTR". It was hoped that by learning about these aggergates with immunofluorescence and transmission electron microscopy using immunogold labeling, they could shed light on what occurs in cystic fibrosis cells expressing misfolded ΔF508-CFTR (85).

CFTR produced in Sf9 insect cells in the baculovirus system has always necessitated the use of the strong detergent, SDS, to solubilize CFTR from the insect cell membranes. In 1998, Scarborough et al. used the synthetic detergent LPG in place of SDS during solublization of the CFTR from insect cell membranes. Not only did LPG detergent solubilize 100% of the CFTR from the insect cell membranes, it also kept underglycosylated CFTR from aggregating during subsequent steps of purification (74). Because of this finding, as well as a need to switch to nickel-column chromatography for purification, we incorporated LPG in place of SDS into our purification protocols in the spring of 2001.

Short synthetic peptides based on transmembrane sequences of ion channels have been shown to possess properties similar to ion channels in regards to both ion conductance and ion selectivity. In 1998, Wigley et al. from the Howard Hughes

Medical Institute synthesized peptides based on the transmembrane helices 1 thru 6 of the N-terminal half of CFTR and reconstituted these into liposomes using SDS detergent (86). They then performed FTIR and CD spectroscopy to determine that all 6 CFTR transmembrane-based peptides were able to form predominantly alpha-helical structures in the membranes of the liposomes as has been predicted by hydropathy analysis for CFTR as early as 1989. In addition, peptide 6 was shown to undergo a shift from the alpha-helical to the beta-sheet form when dissolved in 20% methanol. This study complemented previous mutagenesis studies indicating that helix 6 is important for the formation of the CFTR pore region (87). Other researchers have used peptides based on CFTR primary structure. Peptides mimicking transmembrane sequences of CFTR have been designed for the purpose of determining if specific residues have inherent antioxidant activity within the cell membrane. In 2000, Moosmann and Behl reported that tyrosine and tryptophan-containing transmembrane peptides based on the sequence of CFTR (but not phenylalanine derivatives) prevented oxidative lysis in clonal and primary cell cultures, possibly by inhibiting membrane lipid peroxidation. These investigators speculated that certain nerve disorders such as Alzeheimers disease could be caused by a lack of these amino acid-containing proteins in the predominantly low-protein, high-lipid contents of neuronal cell membranes (88). Six years earlier, Oblatt-Montal et al. from the University of California San Diego had synthesized peptides based on the first 6 transmembrane helices of CFTR in order to determine which may be involved in forming the pore structure. Upon incorporation of these CFTR-based transmembrane peptides into lipid bilayers, it was found that only peptides mimicking endogenous transmembrane

helices 2 and 6, when mixed together and co-reconstituted into liposomes, were capable of forming anion-selective currents (currents which were 95% selective for anions over cations) similar to selectivity of wild-type CFTR, and the conductance rate also was nearly identical with that reported for native CFTR (89).

Because of a lack of structural information about the NBD domains of CFTR, in 2000, the NBD1 domain was expressed as a soluble peptide with a poly-histidine tag, allowing expression of NBD1 in *E. coli* but with the added benefit that the NBD domain didn't need to be renatured after recovery form the bacteria. NBD1 structure determination was then attempted by Duffieux et al. using N15-H1 2D NMR. While they did not succeed in determining the structure of the NBD1 domain, they did report that the domain they expressed in E. coli was folded and would likely at some point yield a suitable structure (90).

Also in 2000, Gerceker et al. from Harvard Medical School made a GFP-CFTR fusion protein in MDCK (Madin-Darby canine cells) epithelial cells to quantitate the ability of CFTR to serve as a receptor in the intestines for the endocytosis of the *Salmonella typhi* bacterium that causes typhoid fever. Using both flow cytometry and confocal laser microscopy, it was determined that there was a strong correlation between intermediate GFP-CFTR expression by these cells and their infection by *S. typhi* (122). It had previously been shown that CFTR could indeed act as a receptor for *Salmonella typhi* in gastrointestinal epithelial cells for submucosal translocation, providing a possible explanation for the theorized "heterozygous advantage" of carriers with a single ΔF508 mutation (123).

The cDNA for CFTR has been transfected into a large number of different cell lines using diverse methods, including direct injection into the nucleus, calcium phosphate precipitation, cationic liposome endocytosis, polylysine peptide conjugation (91), electroporation, and virus transfection. In 1999, Truong-Le et al. from Johns Hopkins University transfected the CFTR cDNA into the cell line 9HTE, a human epithelial cell defective in chloride transport, using calcium and gelatin nanospheres with a covalently bound transferrin protein. The transferrin apparently was able to serve as a ligand for the transferrin receptor (92) on the surface of these 9HTE cells, binding of which was then followed by endocytosis of the DNA conjugate. Their transfection and protein expression efficiency was reported at almost 50%.

CFTR is glycosylated on two asparagines in the fourth extracellular loop, however a comprehensive study attempting to characterize the exact sequences of carbohydrate was not performed on exogenously or endogenously expressed CFTR until 2000. Using newly developed FACE (fluorophore-assisted carbohydrate electrophoresis), O'Riordan et al. found that purified CFTR from CHO cells contained polylactosaminoglycan sequences as did endogenously produced CFTR in human T84 cells. However, CFTR produced in Sf9 insect cells only received oligomannosidic saccharides with fucosylation on the innermost GlcNac (93). Fortunately, differences in glycosylation of CFTR have not been shown to affect its chloride channel activity significantly (94).

By the late 1990s, it was becoming clear from mutagenesis studies that the NBD domains and the R-domain of CFTR somehow directly interact, but proof of this direct interaction was still lacking. Lu and Pederson (from JHU) produced in *E. coli* (as

soluble proteins), the NBD1 domain of CFTR fused to the R-domain of CFTR (NBD1+R), while the NBD2 domain was made as a separate soluble protein. They then were able to detect a direct physical interaction using various assays. In 2000 they detailed how direct physical interactions were found to occur among these domains using four separate protein-protein interaction detection methods. First, they labeled the NBD2 domain polypeptide at the C-terminal end with a fluorophore (CPM), whereby they next detected a red-shift in the lambda max of the attached fluorophore upon addition of unlabeled NBD1+R protein. In their second method, gel filtration chromatography was used to find that the NBD1+R domain, when mixed with the NBD2 domain, came off the size column sooner than either one alone, providing further evidence that these domains had interacted and formed a larger molecular weight structure, as it was found in the excluded volume. In their third experiment, they employed native-PAGE and determined that, when both proteins were allowed to interact, they moved with an R_f value between that of either protein alone, also signifying that an interaction was taking place. Lastly, they performed two trypsin digestions, one on both a mixture of the domains together, and either domain alone. Digestion of the mixture was found to occur at a slower rate than either domain by itself, which was also taken as an indication that they had detected an interaction between these NBD and R domain polypeptides (95).

In spite of the large amount of progress that has been achieved with CFTR, many questions remain unanswered. It is still not proven, for example, which regions of CFTR constitute the pore structure. Structural knowledge at this level of resolution has

the potential to aid in the development of drug treatments designed to increase conduction of mutant CFTR in the CF airway, thereby providing for a possible new treatment direction. Also absent is a structure at atomic resolution for any portion of CFTR. It is believed that the NBD1 domain may soon be solved to atomic resolution. This is significant in large part because the vast majority of naturally occurring mutations that cause CF take place in the NBD domains. It would be useful as well to have a more thorough understanding of the way in which gating of CFTR occurs. Arguably, gating is the most important, and yet least understood aspect of ion channels. Attempts at understanding CFTR channel gating has led to conflicting and often confusing results. For example, some investigators have stated that they believe ATP hydrolysis at the NBD domains is a necessary step for channel opening, while more recent evidence seems to suggest that specific substates can exist whereby hydrolysis does not necessarily have to occur in order for the channel to open, albeit to a lower conductance state (28, 96). There are also physiological questions concerning the role CFTR itself plays in progress of the disease. CF is a disease occurring over a relatively long period of time (decades), indicative of complications beyond faulty chloride conduction. One question still unanswered is how a defect in a single chloride channel (which is not expressed in large amounts) can have such far-reaching effects in the lung and elsewhere in the body. It now appears that some of these answers will come from future studies involving the now well-established ability of CFTR to regulate several other ion channels involved in mucus hydration. This list of ion channels CFTR is able to regulate now includes the outwardly rectifying chloride channel ORCC (97, 98), the amilioride-sensitive sodium

channel ENaC (99), the chloride-bicarbonate exchanger Cl-/HCO3- (100), and the ROMK potassium channel (101).

CHAPTER 3

ENGINEERING A POLY-HISTIDINE TAG ONTO CFTR TO AID IN PURIFICATION

Introduction

A recombinant baculovirus carrying the cDNA for wild-type CFTR has been available in our laboratory since 1994, however the wild-type CFTR protein it coded for had the drawback that purification necessitated use of hydroxyapatite chromatography. This presented a serious problem when it was necessary to obtain CFTR in milligram amounts. While CFTR protein in microgram amounts was suitable for atomic force microscopy experiments (chapter 4), more protein was required for extensive cryo-electron microscopy studies planned at UCSD. The purification procedure calling for a hydroxyapatite chromatography step was found to be unpredictable and non-robust when scaling up production of CFTR protein; both as far as its quantity and purity of CFTR. Because of this lack of reproducibility (in our hands) using hydroxyapatite chromatography, it was decided in the Summer of 2000 to engineer a 10 histidine tail onto the C-terminal end of CFTR to aid in a faster, more efficient and reproducible purification using nickel-affinity chromatography in place of the less-specific and more subjective hydroxyapatite chromatography procedure.

Background

Use of the baculovirus expression system has grown steadily following the first successful non-viral protein production of β-interferon in an insect cell line using a baculovirus in 1983 by Smith et al. (102). Some reasons for this trend towards insect-

produced protein include authentic post-translational modifications characteristic of eukaryotic cells (such as glycosylation, phosphorylation, acylation), use of cells which are grown at room temperature and require no carbon dioxide, the ability of insect cells to properly fold and target large transmembrane proteins, and relatively high amounts of active recombinant protein are often recovered using gentle non-denaturing conditions. In fact, up to 50% of total cellular protein in some cases has been reported for certain soluble foreign proteins (103). Generally, to produce transgenic baculoviruses, the endogenous viral gene encoding the polyhedrin structural protein normally expressed in large amounts late in infection, is replaced during a homologous recombination event in the insect cell nucleus with a transgene of interest which has been previously spliced into a plasmid known as a "transfer vector". The baculovirus most often used for foreign protein production is the *Autographica californica Nuclear Polyhydrosis Virus* (AcNPV), a 117 KB double-stranded DNA virus, able to infect at least 3 different cell types *in vitro*, including Sf9 cells. The Sf9 insect cell line was developed in 1977 from a *Lepadoptran* species of caterpillar, and can be grown either semi-adherently or detached in spinner flasks (104).

The exact variation of expression system used in this study differs in some significant ways from the original procedure developed by Smith et al. in 1984 in that homologus crossover not only exchanges the very late-expressed polyhedrin gene from the virus with the CFTR cDNA from the transfer vector, but at the same time reconstitutes the activity of an essential open reading frame, ORF 1629, a critical part of which was previously removed during linerization of the viral genome using a rare-cutting restriction enzyme

(Invitrogen Inc). This prior linerization of viral DNA also improves recombination frequency with the transgene, as well as making it all but impossible for self-circularization of virus to take place without the desired insert; recircularization which would otherwise lead to a higher background and therefore make selection of a suitable recombinant virus clone that much more difficult (Invitrogen, Inc). In addition to linerization of viral genomic DNA, a portion of the *E. coli* B-galactosidase gene was engineered into this transfer plasmid to make selection easier upon recombination with the transgene.

Modifications to the primary structure of CFTR have been reported in the literature periodically, and many of these have not appeared to harm chloride channel activity of CFTR in any measurable way. Epitopic FLAG sequences (105), glycosylation signal sequences (63), and viral antigenic sequences (75) have been substituted into several of the loops of CFTR without apparent detriment. In 1997, Ramjeesingh et al. engineered a poly-histidine tag onto the C-terminal end of CFTR to aid in purification (73). This work was repeated by a group at the University of North Carolina, Chappel Hill who, in 1998 placed the same poly-histidine tag onto the C-terminal end of CFTR, except in this instance using a yeast *S. cerevisiae* protein expression system as opposed to the insect cell-based procedure (74). In 1998, Dartmouth researchers added an N-terminal green fluorescence protein construct (GFP-CFTR) onto CFTR for the purposes of tracking CFTR intracellularly in mammalian cells. Importantly, they found no change in chloride channel properties when measured with whole cell patch-clampimg (121).

Overview of poly-histidine tagging of CFTR

Initially, the 3' end of CFTR was amplified using PCR employing a reverse primer designed to incorporate 10 histidine codons, followed by a stop codon and finally a *KpnI* restriction site. Conversely, the forward primer was designed to hybridize to an upstream, interior portion of the CFTR gene where there already existed a natural *NcoI* restriction site within the coding sequence. This resulting amplified 3' fragment was then digested with the same two restriction enzymes used to cut out the original 3' end of the wild-type CFTR gene. The new, PCR-generated 3' fragment of CFTR with the poly-histidine codons near the end was then ligated back into the original CFTR cDNA which had previously been placed into a more recent version of a baculovirus transfer vector (discussed above). The transfer vector was then co-transfected along with linearized viral DNA into cultured insect cells. Later, the resulting recombinant baculovirus containing the CFTR cDNA was selected for using blue/white colony screening following infection of insect cells at various dilutions of recombinant virus. Lastly, recombinant viral stocks were obtained using this single viral clone, and large-scale stocks of insect cells infected with them and tested for poly-histidine-tagged CFTR using nickel chromatography purification. This was followed by silver staining and western blotting of the SDS-PAGE gels to verify amount and purity of CFTR obtained.

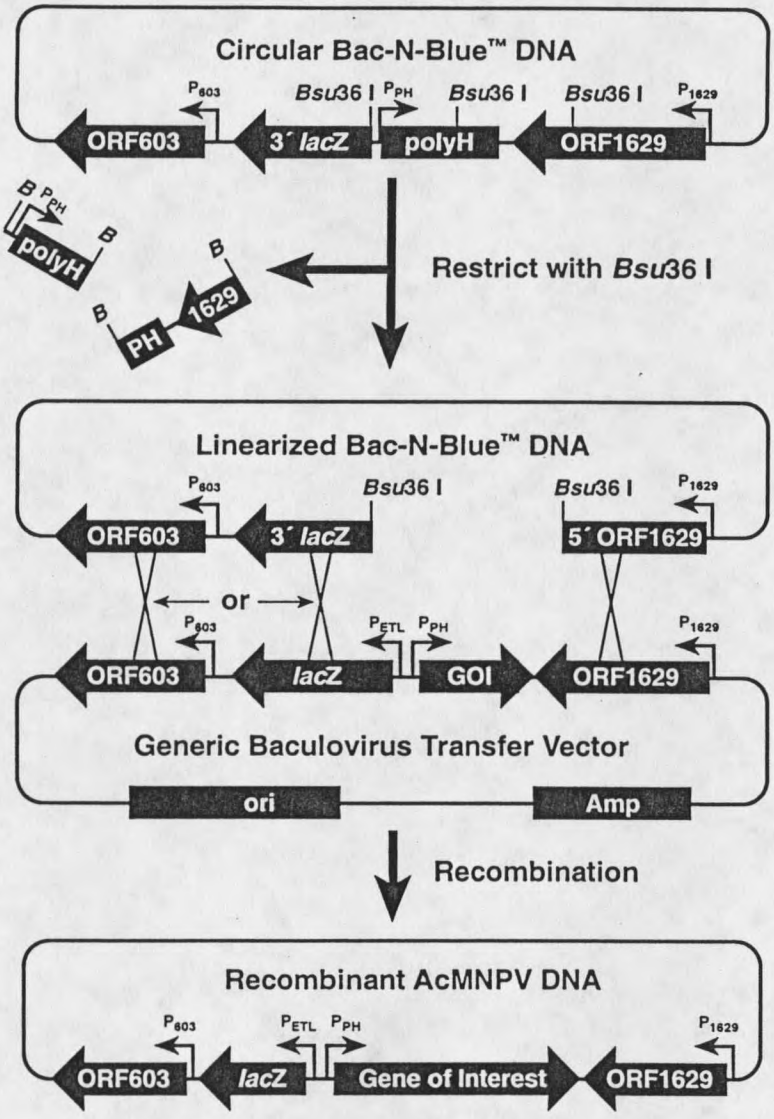

Figure 6: **Cross-over of GOI** (i.e. CFTR cDNA) from transfer vector into baculovirus genome. Note that first step (BsuI digestion) was performed by supplier (Invitrogen), and cross-over of CFTR cDNA is accomplished by transfer of missing portion of an ORF necessary for viral replicaton, as well as part of the LacZ gene to aid in detection using plaque assay.

Methods

PCR Amplification of 3' terminus of CFTR

Approximately the last 500 base pairs of the cDNA of CFTR was amplified for the purpose of adding a poly-histidine tag onto the C-terminal end. Two primers were used for this purpose: the forward primer was designed to hybridize immediately upstream to where an endogenous *NcoI* restriction site exists in the CFTR cDNA. The forward primer sequence was: 5'-GCTGTGTCCTAAGCCCATGG-3'. The downstream, reverse primer hybridized at the 3' UTR immediately past the endogenous CFTR stop codon, and contained 10 histidine codons followed by a new stop codon, and finally a restriction site for KpnI at the extreme 5' end of the reverse primer sequence. The reverse primer sequence was as follows:

5'CGGGTACCTCTAATGATGGTGATGATGGTGGTGATGGTGATGAAGCCTTGT ATCTTGCACCTC3' (Genosys Inc). For the PCR reaction: 100 ng pCFTR plasmid template containing the cDNA of wild-type CFTR was combined with 100 pmol of each primer. dNTPs at a final concentration of 0.2 mM were also added. The temperature was cycled 30 times using a thermocycler. The temperature cycle consisted of 95 °C for 0.5 min, 65 °C for 0.5 min, and 72 °C for1 min. The resulting amplified fragment migrated at approximately 500 bp as expected (Figure 7). This contrasted with the corresponding 3' fragment removed from the original CFTR cDNA using restriction digestion, which migrated at approximately 600 bp. This discrepancy in size is attributed to the fact that the *KpnI* site in the wild-type CFTR cDNA was ~100 bp further downstream of the original stop codon and therefore located further into the 3' UTR of

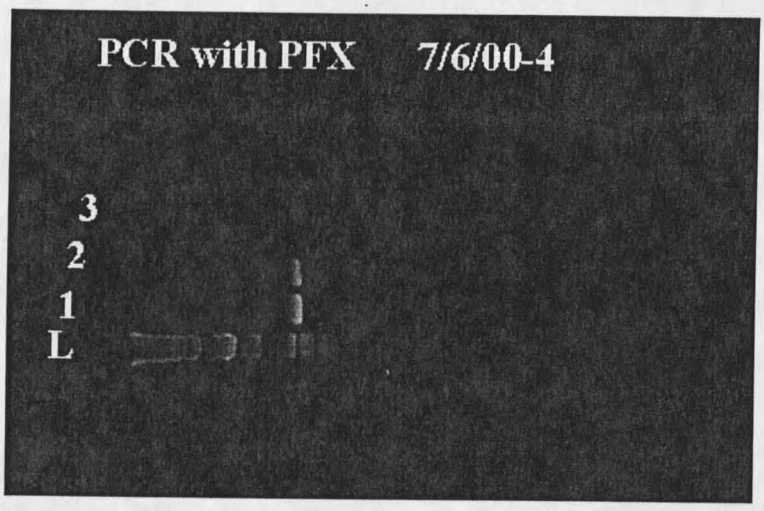

Figure 7: 3′ PCR-amplified fragment (lanes 1 & 2) of CFTR cDNA

the CFTR cDNA. This difference in size between the 3' end fragments of the CFTR gene later enabled verification of the correct placement of the PCR product into the 3' end of the CFTR gene in the new transfer vector simply by double-digesting the finished baculovirus transfer plasmid using *KpnI/NcoI* followed by agarose gel electrophoresis (0.8%). The thermophilic DNA polymerase *Pfu* (Stratagene Inc) was used in the PCR reaction as *Pfu* is generally considered to have a higher copying fidelity than *Taq* DNA polymerase.

Restriction Digestion of New 3' Insert

Because both primers used in the PCR step contained a restriction site, *KpnI* in reverse primer and *NcoI* in forward primer, it was possible to create sticky ends using these two restriction enzymes by performing a *KpnI/NcoI* double-digest of this PCR-amplified 3' insert. It was not possible, however, to determine efficiency of the double-digestion except by gauging the success or failure of ligation with the original 5' portion of the wild-type CFTR cDNA in the following steps. Approximately 10 μg of the PCR-amplified 3' insert was double-digested with *KpnI* and *NcoI*. These restriction enzymes were later heat-inactivated at 60 °C following a 1 hour digestion at 37 °C. This double-digested insert was then precipitated with ethanol, and the concentration of the recovered, concentrated 3' insert estimated using agarose gel electrophoresis to be approximately 500 ng/μl. (results not shown)

Restriction Digestion of pCFTR and pBlueBac with *PstI*

Both the new pBlueBac transfer vector plasmid and the original pCFTR plasmid carrying the wild-type CFTR cDNA were digested with *PstI* for 1 hour at 37 °C (Figure 8, lanes 2 and 4 respectively). Afterwards, *PstI* was heat-inactivated at 65 °C for 10 minutes. Next, calf intestinal phosphatase (CIP, Promega Inc) was used to dephosphorylate pBlueBac 3' sticky ends in order to discourage self-ligation of the plasmid not containing the insert. The entire wild-type CFTR cDNA insert was cut out of an agarose gel (0.8%) and purified using ethanol precipitation (Figure 8, lane 4). Finally, cut and dephosphorylated pBlueBac was run on an agarose gel, cut from the gel,

and also concentrated with ethanol (results not shown). Using the ethidium bromide-stained agarose gel, the efficiency of the *PstI* digestion was estimated at around 75 % by comparing the intensity of the bands with that of the ladder.

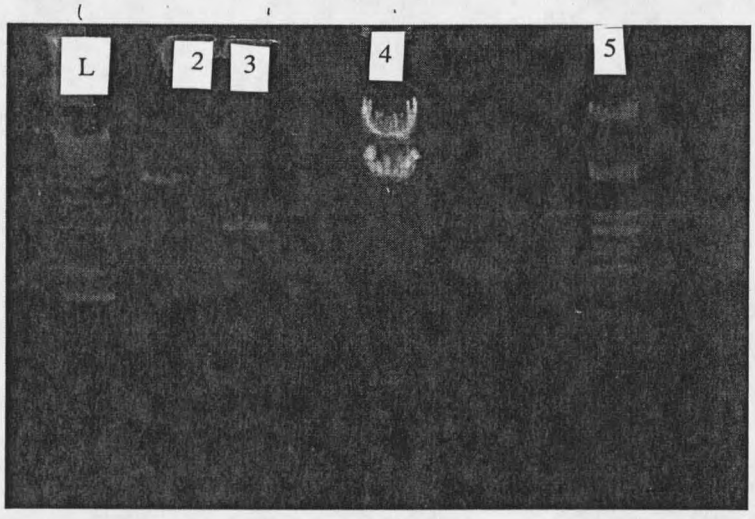

Figure 8: Agarose gel (0.8%) EtBr-stained. Lane 2: pBlueBac cut with *PstI*. Lane 3: uncut supercoiled pBlueBac. Lane 4: pCFTR with CFTR cDNA cut out (lower band) with *PstI*. Lane 5: λ-phage genome cut with *PstI*.

Ligation of pBlueBac with Original CFTR cDNA

1 µl T4 DNA Ligase (Promega) was used for each ligation. Ligation was performed at 16 °C for 14 hours. 2 µl wild-type CFTR insert plus 1 µl pBlueBac transfer plasmid constituted the first ligation. The second ligation was exactly the same, but without CFTR insert and was therefore used as a negative control in order to assess degree of self-ligation of the plasmid alone without insert.

Transformation of New Transfer Vector (pBlueBac)

Containing Wild-type CFTR Insert

Chemically competent DH5α bacterial cells (GibcoBRL Inc) were heat-shocked in order to transfect these cells with ligated plasmids. Negative control was simply plasmid only, with no insert present. The result was that 2 colonies on the negative control LB/Amp100 plate and 12 on the plate with insert ligation. All 12 colonies were expanded in liquid LB media with ampicillin (100ng/ml). Plasmids were purified from DH5α cells using PERFECTprep Plasmid DNA kit® (5 prime-3 prime Inc) . Yield was estimated by agarose gel electrophoresis (0.8%) at approximately 500 ng/μl for each plasmid clone (results not shown).

Ascertaining Direction of Insert in Plasmid

Because the only restriction enzyme used was *PstI*, it was necessary to determine the direction of insert in plasmid, as both orientations were equally likely. *XbaI* digestion of 2 μg each plasmid clone enabled confirmation of the orientation, as only CFTR insert in correct orientation will yield an *XbaI* fragment of 650 bp in length. Of the 12 plasmids, only 3 total had insert. Clones 1 and 3 contained the wild-type CFTR insert in backwards and clone 9 had it in forwards (results not shown). Clone 9 was therefore used for the rest of the procedure.

Digestion of Clone 9 with *KpnI/NcoI* to Remove 3' end of Original CFTR gene

The two restriction enzymes *KpnI* and *NcoI* were used to digest out and remove the 3' wild-type insert. The resulting product was ascertained on an agarose gel (0.8%) for digestion efficiency. This double-digestion was performed in order to remove the portion of the CFTR terminus which would ultimately be replaced by the PCR-amplified product with the poly-histidine codons incorporated. The results were that most of the plasmid had been double-digested successfully and the appearance of a ~600 bp fragment meant that the original 3' end had been removed as intended (results not shown). Digested Clone 9 was then cut from the gel using a razor blade and purified using ethanol precipitation. The purified Clone 9 transfer plasmid carrying only the 5' beginning of the CFTR wild-type insert would be used in ligation with the PCR-amplified insert in the following step.

Ligation of Modified CFTR 3' Insert with Baculovirus Transfer Plasmid Containing 5' CFTR Insert

As both insert and plasmid were double-digested with the same restriction enzymes, they both had complementary sticky ends and could be ligated together directly. Approximately 500 ng (1 µl) of each were added along with T4 DNA ligase and incubated at 16 °C overnight. A ligation without the presence of CFTR 3' insert was also done as a negative control and to assess efficiency of plasmid digestion and dephosphorylation.

Transformation of Ligation Products into *E. coli*

A single microliter of each of the above ligation products was added to DH5α bacteria using the method of heat-shock (as above). Cells were then spread onto LB/Amp100 agar plates and incubated overnight at 37 °C. Only 2 colonies were found on the control plate, however 7 were on the plate containing the ligation with 3' insert. These 7 colonies were then expanded in liquid LB/Amp100 media and plasmid was purified from each clone using PERFECTprep Plasmid DNA kit® spin columns.

Verification of Presence of New 3' Insert in Transfer Vector

A double-digestion of the above positive clones was performed using the *KpnI/NcoI* restriction enzymes. As the correct insert containing the poly-histidine codons should produce a fragment just below the 500 bp ladder fragment, this test was straightforward and was successful. Agarose gel electrophoresis clearly showed a 500 bp fragment and it was determined to go forward at this time with 6 clones showing positive responses (results not shown).

Sequence Verification of Cloned Insert Created with PCR

Because PCR is known to produce a high number of mutations when compared to endogenous replication of DNA in intact organisms, it was necessary to determine the sequence of the 3' insert containing the poly-histidine tag. Approximately 100 ng of

plasmid and 3.2 pmol forward primer (same primer used for PCR earlier) were used for sequencing. The sequencing method was based on the dideoxynucleotide chain termination method pioneered by Sanger. It was also necessary to design a second primer in order to obtain sequence verification immediately downstream of where the forward primer in the first sequencing reaction had hybridized, as approximately 50 bases downstream of it were unintelligible. This second, reverse sequencing primer hybridized at base 4357 of the CFTR cDNA. The sequence of the primer was: 5'-CCAGCATGCTCTATCTG-3'. Following success of both sequencing attempts, it was discovered that of the 6 clones tested, only a single insert, clone 6, had no errors in the 3' PCR-created sequence. This large number of mistakes in the other clones was somewhat surprising as the PCR polymerase enzyme used, *Pfu*, is generally considered a high fidelity enzyme.

Co-Transfection of Baculovirus DNA with Transfer Plasmid into Insect Cells

Liposomes have become the established method of choice for the introduction of foreign DNA into insect cells for the purpose of creating recombinant baculoviruses. Liposomes were supplied by Invitrogen. Sf9 insect cells were grown to approximately 60% confluency in T-25 flasks and were therefore in log phase. 4 μg of clone 6 transfer plasmid containing the CFTR cDNA was added to 10 μl linearized baculoviral genomic DNA. Then, 1 ml of serum-less TNM-FH media was added. Next, 20 μl of liposomes were added to this tube and the entire mixture briefly and gently vortexed. This mixture

was then allowed to incubate at room temperature for 15 minutes. At this point, sf9 cells were subjected to the transfection mixture (drop-by-drop) and the flask was rocked in a side-to-side manner at 27 °C. After 1 hour, 4 ml fresh serum-containing media was added to the flask and the flask was sealed in a plastic bag and incubated for 3 days at 27 °C. Signs of viral infection, primarily a uniform roundness of cells, became visible under microscope after 3 days. At this time, 4 ml of media from T-25 flask was removed for use later in obtaining a single viral clone using the plaque purification assay.

Selection of Recombinant Baculovirus

Here, uninfected sf9 cells were expanded to near-confluency on sterile 60 mm petri dishes and infected with the above media (containing virus), which was at this point at an unknown concentration (i.e. unknown viral titer). For this reason, virus-containing media was first diluted to several concentrations ranging from 10-10,000 fold dilution before applying to cells on plates. Cells in each plate were infected with 1 ml of this diluted virus (which was at this point still a mixture of recombinant virus and virus without the CFTR cDNA) and rocked side-to-side at 27 °C for 1 hour. After this, a specific low-temperature melting agarose (Invitrogen) dissolved at 2.5% in TNM-FH media along with chromogenic substrate (150 μg/ml) X-Gal in DMSO (Fisher) was applied to cells in order to form a layer on the top which would ideally keep any cells (and therefore recombinant virus) from moving around on the plate during the 4 day procedure. Blue plaques were visualized on several of the plates diluted with virus at the 1:100 and 1:1000 dilution. As the blue colonies were spaced farther apart on these

plates, several colonies were taken as agarose plugs with a plastic transfer pipette, removed and placed individually into T-12 flasks of near-confluent log-phase sf9 cells. This presumably insured that individual viral clones were being isolated away from virus not containing the CFTR insert and then grown in T-12 flasks in high enough amounts to begin making large-scale stocks of virus to begin expression of CFTR protein.

Growing Viral Stocks for Expression Trials

Once the above T-12 flasks of cells started showing signs of infection (indicated by a uniform swelling and later detachment of cells from flask bottom), several drops of media were removed from one of these and added to a 100 ml spinner flask of Sf9 cells grown to a density of approximately 0.75×10^6 cells/ml. After 3 days, Sf9 cells were showing obvious signs of infection, including swelling and lysing of cells, and therefore were harvested for virus. This was accomplished by spinning cells in a centrifuge at ~1000g for 5 minutes. Virus-containing supernatant was decanted and designated as "P-1 stock" and later used to induce expression of CFTR protein. Baculovirus is a large DNA virus capable of light-scattering which resulted in the supernatant appearing slightly cloudy. This was therefore taken as an indication that baculovirus was probably present in sufficiently high amounts for protein production.

CFTR Expression

After 100 ml of baculoviral stock was obtained by expansion (above), two T-75 flasks were grown nearly confluent with sf9 cells for expression of protein. After removal of

media, 5 ml of virus stock was added to each flask and allowed to incubate on cells at room temperature for 1 hour. Then baculovirus was removed and 15 ml new media containing serum was placed onto cells and cells allowed to incubate at 27 °C for 48 hours, sufficient time for CFTR protein (under control of the very late polyhedrin promoter of the baculovirus) to be made. Cells appeared virus-infected after only 24 hours. After a total of 48 hours, infected cells were harvested by gently pipetting cells off flasks. Then, these cells were centrifuged at 1000g for 10 minutes. Cell pellet was washed 1X with PBS buffer and spun again as above. PBS supernatant was removed and pellet was kept to assess protein production.

Histidine-tagged CFTR Protein Purification

Cells from above were lysed by solubilizing in alkaline extraction solution (10 ml, pH 10.1) to which protease inhibitors (leupeptin, aprotinin, PMSF, 50 μg/ml each final concentration) had been added. Resulting cell membranes were pelleted by ultracentrifugation (35K, 40 min). Membranes were then solubilized with 2 ml of Buffer A (10 mM Tris, pH 7.9) containing 2% lysophosphadidyl glycerol (LPG, Avanti Lipids) and protease inhibitors. Nickel chromatography was performed as a slurry initially (Novagen Inc). One milliliter of suspended beads were charged with nickel sulfate according to instructions from Novagen. Solubilized total membrane protein was added to slurry of nickel-charged beads and rotated in a tumbling motion for 5 hours at 4 °C. After 5 hours, the slurry was transferred to a column and allowed to settle. The first wash was performed with binding buffer (5 mM imidazole in Buffer B: 15 mM Tris,

500 mM NaCl, 0.02 % LPG, pH 7.9). After rinsing with 7 ml binding buffer, 2 ml of this was saved and concentrated (centricon 50, Amicon Inc) for testing on SDS-PAGE and western blot. This was repeated for the 3 more steps (60 mM imidazole in buffer B) (430 mM imidazole in buffer B) (and 1000mM imidazole in buffer B). All 3 samples were tested for presence of CFTR (using SDS-PAGE/silver stain and western blot).

CFTR Detection

Above samples were concentrated to approximately 250 µl each using 50 KD cut-off centrifugal filtration (Centricon, Amicon Inc) and prepared for SDS-PAGE analysis using 6X SDS sample buffer (250 µl each sample plus 60 µl 6X sample buffer). Samples were then heated for 20 minutes at 40 °C as CFTR aggregates above this temperature. Two 7% acrylamide gels were prepared each consisting of a stacking layer. Detergent-solubilized membrane protein samples were loaded individually (20 µl for silver stained gel and 30 µl for western blot gel) into each lane. Following the electrophoresis run (200 V, 40 min), one of the gels was silver-stained (ICN Biomedicals Inc) and the remaining gel subjected to western blot procedure. PVDF membrane (Schleicher & Schuell) was used for the transfer and subsequent blotting. Silver-staining was done using a Rapid-Ag-Stain™ kit (ICN Biomedicals Inc) which has a protein detection limit of approximately 50-100 ng. Western blotting procedure utilized a primary monoclonal antibody directed against the R-domain of CFTR (Genzyme Inc). Secondary antibody was an anti-mouse goat antibody conjugated with alkaline phosphatase (Sigma

Inc). Detection of CFTR protein therefore was via a colorimetric change, in this case purple, on the PVDF membrane (Westran, Schleicher & Schuell).

Results

Silver-stained SDS-PAGE results are shown in figure 9. CFTR migrated at ~130KD as expected. Lanes 1-5 are fractions from the Nickel column (Novagen) eluted with increasing amounts of imidazole. Lane 1 is protein removed with 5 mM imidazole. This is presumably protein which has either not stuck to the nickel at all, or is bound so loosely that only 5 mM imidazole is able to knock it off into the eluent. Lane 2 was protein eluted with 60 mM imidazole. Here, there is an increase in both CFTR as well as contaminant. Lanes 3 and 4 are CFTR eluted with 450 mM imidazole in the buffer.

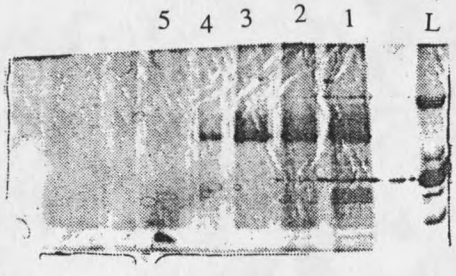

Figure 9: **Silver-stained SDS-PAGE gel** of protein fractions from nickel affinity purification. Lane L: Molecular weight ladder (top to bottom: 150 KD, 120 KD, 95 KD, 85 KD). Lane 1: Protein removed from column in 5 mM imidazole wash. Lane 2: Protein (i.e. pure CFTR) from 60 mM imidazole wash. Lane 3: CFTR eluted in 450 mM imidaole. Lane 4: spill-over from lane 3. Lane 5: results of 1000 mM wash with imidazole.

Here, CFTR is essentially pure. Note that lane 4 is simply spill-over from what was loaded into lane 3. Lane 5 was an elution with 1000 mM imidazole. There was no protein at all in this fraction.

Results of western blot are shown in figure 10. Here, the procedure was performed on the same samples as silver stain (above) with the addition of a sample of the membrane vesicles before purification in order to ascertain efficiency of LPG detergent in solubilizing CFTR from the membranes (lane 1). Lane 1 is solubilized total protein from membranes prior to purification. In addition to CFTR, low molecular weight contaminant can be seen which also stained with the antibody. Lane 2 is of total protein that flowed through the column without ever binding. This is similar to what appears in lane 1. Lane 3 is CFTR from 5 mM imidazole step. Here, no proteolysis is apparent. Lane 4 is CFTR from 60 mM imidazole step, and Lane 5 is CFTR from 450 mM step.

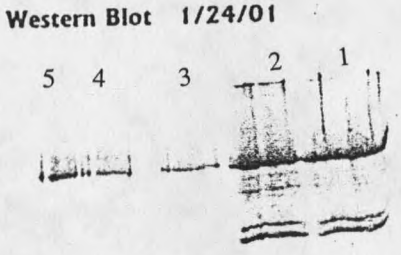

Figure 10: **Western Blot of CFTR** with mAb to R-domain. Lane 1: membrane insoluble CFTR. Two bands below are partially degraded CFTR. Lane 2: flow thru of CFTR that did not stick to nickel-affinity column. Lane 3: CFTR eluted from column in 5 mM imidazole. Lane 4: CFTR eluted in 60 mM imidazole. Lane 5: CFTR eluted in 450 mM imidazole.

Discussion

The results of the silver staining procedure clearly showed that imidazole in increasing amounts knocks CFTR in purer form from the nickel column. As a precaution, a 1M imidazole step was performed to ascertain the degree of CFTR still stuck to the column after the 450 mM imidazole step (lane 5). There was no CFTR found in the 1M imidazole-containing fraction. The western blot revealed that much degredation from proteolysis had taken place in both the flow-thru and the membrane solubilized, unpurified CFTR. This was also not unexpected, as the proteases were not purified away from the CFTR in these fractions.

It appears from these results that CFTR-10-His either has an inherently low affinity for the nickel or perhaps there are not enough beads. Evidence for low affinity are that very little CFTR sticks to the column and is later eluted when solubilized protein is simply run through the column without the benefit of being as a slurry first (results not shown). However, it is also true that when the amount of nickel beads in the slurry are doubled, the amount of CFTR gained is also approximately doubled, leading to the conclusion that there are not enough prime binding sites. In the future, it is perhaps worthwhile to use excess beads as this has shown each time to increase the amount of CFTR gained during the purification (results not shown).

Proteases have been a continuing problem when using the baculovirus system to express CFTR. Three different findings that have been noted during purifications over the last several years have lead to this conclusion. The first is that the same two low molecular weight bands reappear each time there has been contamination, and both of

these bands stain at the same apparent molecular weight in the western blot (which uses an antibody to the R-domain, an antibody which has never been known in our lab to cross react with other proteins). The second reason is that each time there is a decrease in the yield of CFTR, there is a corresponding increase in the intensity of the two low molecular weight bands. A third reason to believe proteases have been the primary problem is that whenever inhibitors are not used fresh (or occasionally when some have been omitted), there tends to be a corresponding increase in the amount of CFTR loss. Baculovirus genomes code for an endogenous cysteine protease and this could be a source of protease contamination during purification. It is also possible that Sf9 cells are producing the proteases as well, as these proteases are equally likely to make their way into the purificaton protocol.

The western blot results (figure 10) suggest that much of the CFTR is not being solubilized by the detergent LPG and is being lost in the unsolubilized membrane fraction. This sample was taken from what is usually a relatively small amount of membrane pellet that is discarded following the final spin in the ultracentrifuge before (i.e. just adding solubilized protein to the slurry). Even though it appears to be a large amount of loss, it should also be noted that there was only a total volume of approximately 30 microliters of membrane fraction. When this small volume is taken into account and compared with the several hundred microliter volumes of the eluent fractions, it doesn't amount to nearly as much loss as appears to be occurring like the western blot results suggest.

At one point, CFTR-10-His in our lab was subjected to hydroxyapatite to in part ascertain hydroxyapatite's ability to function as a polishing step (results not shown). The CFTR that was eluted from this column was at least 1 to 2 orders of magnitude less than had been recovered by the nickel-affinity column, further strengthening our belief that this procedure is much more reliable in terms of yield than hydroxyapatite. We also, in 1999, were the first to produce CFTR in the *Drosophila* protein expression system (Invitrogen Inc) as a stable integrant in the chromosomes of these cells (unpublished results). CFTR was found using this system upon induction of a metallothionein promoter with copper, but was barely detectable using western blotting (not shown). For this reason, the *Drosophila* expression system was abandoned by us in favor of the bacluovirus expression system for the production of human CFTR protein. In the future, when proteolysis becomes a problem using the nickel-affinity procedure, it may be necessary to polish the CFTR fractions with either density gradient centrifugation or gel exclusion chromatography. Newer proteases could also be tried, as well as thoroughly autoclaving buffers and equipment used for purification.

Conclusion

When no proteolysis takes place, this procedure has the potential to be a one-step purification of CFTR, the only one we are aware of. Other reports in the literature where CFTR-10-His has been nickel-affinity purified indicate that a second polishing step was necessary. In the yeast expression system by Scarbourough et al. (74), it was necessary to perform density gradient centrifugation to clean the sample of impurities. In

the insect cell protocol by Ramjessheh et al. (54) gel filtration was also done after affinity purification.

CFTR Activity Assays

Unlike many soluble enzyme assays, there are far fewer methods to detect activity of ion channels. This is especially true because ion channels are usually not available in large amounts. Four of the most common methods for monitoring ion travel through a protein channel include fluorescence quenching, radioisotopic ion assays, assays using ion-sensitive electrodes, and, by far the most sensitive of all, the electronic patch clamp method. In the current study, all except the ion-sensitive electrode method were attempted, as there are no halide-sensitive electrodes generally available that are sensitive enough to detect halide ion current. Of the three ion channel assays discussed here, the method based on the patch clamp (performed at Harvard University by Cantiello et al. in 1999) was the only method sensitive enough to detect chloride transport through purified and reconstituted CFTR produced in our lab.

A Potential Fluorescence-based Assay for CFTR in Membrane Vesicles

SPQ (6-methoxy-N-(3-sulfopropyl) quinolinium, Molecular Probes Inc) is a halide-sensitive small organic molecule which is fluorescence-quenched (via a collisional mechanism) by 50% for each 10 mM halide ion present (106). Ascertaining CFTR chloride channel activity in membrane vesicles (as opposed to purified CFTR in liposomes) is in theory much easier than ascertaining CFTR activity incorporated into

artificial bilayers. This would therefore enable measurement of the degree of activity of CFTR before embarking on more costly and time-consuming purification and reconstitution procedures. In the strategy described here, CFTR-expressing insect cells were lysed and then loaded with SPQ using osmosis. These fluorophore-loaded membrane vesicles were, later in the experiment, exposed to chloride from the exterior and a decrease in fluorescence indicating chloride transport into the vesicles was monitored.

Methods

In an attempt to assay CFTR activity in membrane vesicles, a single T-75 flask of baculovirus-infected Sf9 cells was first pelleted by centrifugation and then resuspended in hypertonic 500 mM sodium sulfate/10 mM Tris buffer and lysed using a 27G needle. The resulting vesicles were spun in an ultracentrifuge and pelleted (35K, 40 min). The membrane vesicle pellet was then resuspended in hypotonic transport buffer containing 2 mM SPQ. Membrane vesicles were next allowed to incubate (10 min at room temp) in order to take up the SPQ dye. Then this mixture was subjected to gel filtration chromatography (G-50 beads, Sigma) to remove unexcluded SPQ fluorescent dye (i.e. SPQ not in the membrane vesicles).

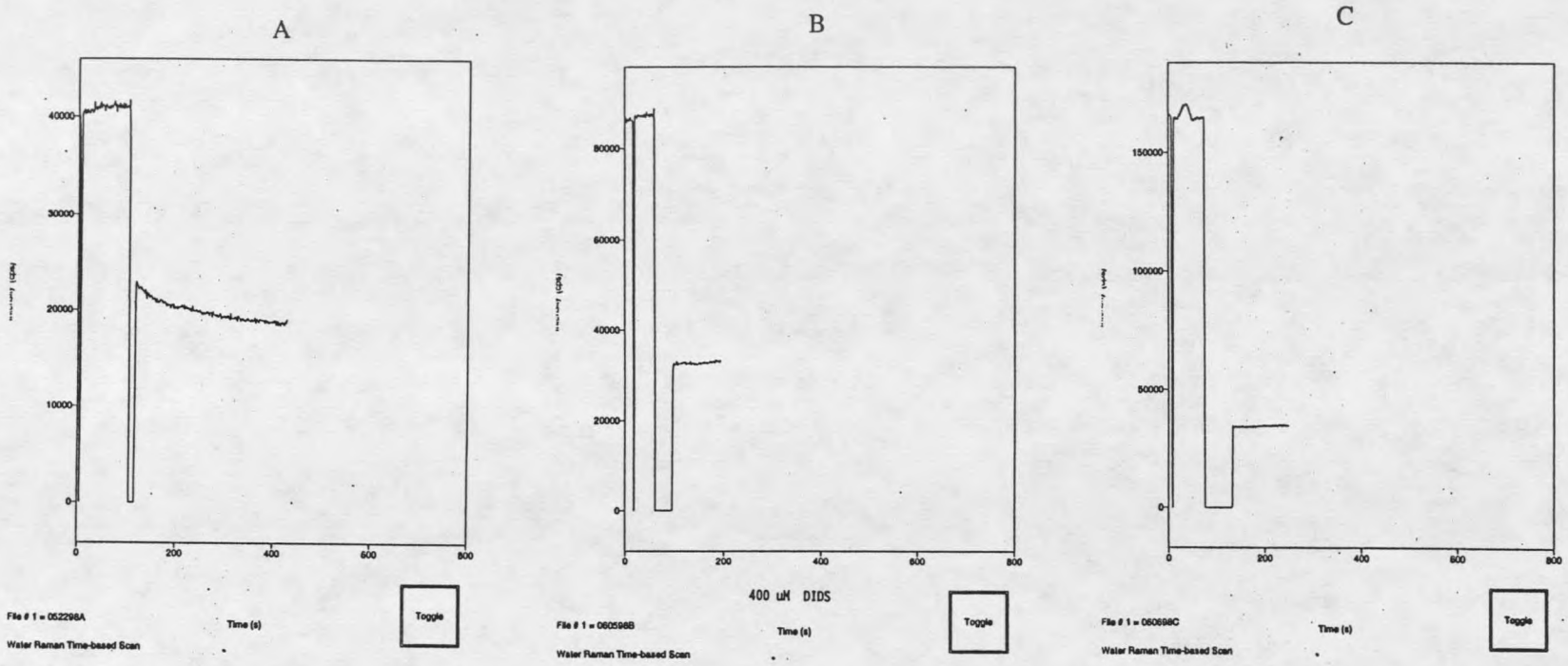

Figure 11: **Fluorescence-based assay for chloride channel activity** of insect cell membrane vesicles. A: addition of chloride at 120 seconds results in a progressive decrease in fluorescence intensity as chloride enters vesicles and quenches SPQ fluorescence. B: DIDS (a broad spectrum chloride channel inhibitor) was added initially. Upon later addition of chloride, no time-dependent decrease in fluorescence as in "A" occurs, indicating that there was no chloride passing through membrane vesicles. C: Same as in "B", except these membrane vesicles were from insect cells infected with baculovirus and were therefore expressing CFTR.

Results

These membrane vesicles were tested for fluorescence after being recovered in the excluded volume. As there was no chloride in the gel filtration buffer at this point, addition of 125 mM KCl (final concentration) to the quartz cuvette was predicted to cause a steady decrease in fluorescence as the chloride makes its way into the membrane vesicles through the chloride channels (Figure 11-A). However, in the presence of the detergent octyl-glucoside, this was not seen, as the integrity of the cell membrane had been destroyed (this loss of membrane integrity was also visible as a decrease in light scattering). When chloride was added to the detergent-solubilized membrane vesicles in the cuvette during measurement of fluorescence, there was therefore no change in fluorescence (results not shown). In another experiment also indicating chloride was indeed flowing through ion channels within the SPQ-loaded membrane vesicles, the general chloride channel inhibitor DIDS (4, 4'-Diisothiocyanato-stilbene-2, 2'-disulfonic acid, Sigma) was added (final concentration: 400 μM) just before the addition of the potassium chloride. As seen in figure 11-B, there was also no decrease in fluorescence visible, suggesting that the vast majority of chloride channels being assayed here were endogenous to the insect cell and were not CFTR (since CFTR is the only known chloride channel *not* inhibited by DIDS when added to the exterior of the channel) (127).

Discussion

Disappointingly, while it was possible using this method to detect endogenous insect cell chloride channel activity, CFTR itself was apparently not detected above background (figure 11-C). One possible reason for this came to light in a later experiment (next section) using an assay for chloride channel activity involving Cl-36. Apparently CFTR, when expressed in insect cells using the baculovirus system, is not produced in high enough amounts to be detectable above background endogenous chloride channel activity when using this fluorescence-based assay.

Whole-Cell Chloride-36 CFTR Assay

Transport of radioactive chloride-36 (figure 12), which is a medium-strength beta emitter, can be useful in the measurment of the rate of chloride transport through chloride ion channels such as CFTR (107). In these experiments, cells both expressing and not expressing CFTR were exposed to Cl-36 for increasing lengths of time, collected and finally counted for radiation using a scintillator.

Methods

In this assay, an 8 well slide was seeded with Sf9 cells grown to near confluency. Next, 4 wells were infected with baculovirus (figure 12-B) for the purpose of producing CFTR protein in these cells. Cells in the remaining 4 wells were left uninfected and used as a negative control (figure 12-A). Thirty-six hours after infection, cells were incubated briefly (15 min) in PBS buffer (pH 7.2) lacking any halide ions except for 1

μCi chloride-36 (ICN Inc.), 5 μl forskolin to activate any CFTR present. Periodically, cells from each well were detached from the slide by gentle pipetting and transferred onto a 0.22 μm filter using a vacuum. These filter-attached cells were then rinsed with PBS buffer lacking radioactive chloride. Filters with the washed cells attached were then placed directly into a scintillation counter tube and counted.

Results

When DIDS, a broad spectrum chloride channel inhibitor, was present on the outside of the vesicles (final concentration of 1 mM), there was an overall decrease in chloride channel activity, however in the wells where CFTR was present in the insect cell membranes the decrease in activity was not as dramatic as with cells not expressing CFTR. In wells with cells expressing activated CFTR, there was a steady increase in chloride-36 accumulation inside the cells, even in the presence of DIDS, which would have shut down all endogenous insect cell chloride channels. It should also be noted that PKA was not necessary in this assay because cells were activated with forskolin (10 μM) for 15 minutes. Forskolin is a membrane-permeable small molecule able to directly activate adenylyl cyclase in cells treated with it.

Discussion

It was concluded that this method was able to detect CFTR in intact insect cells, however the amount of channel activity when compared to endogenous chloride channel activity was disappointingly low, and may explain why detection of CFTR using the

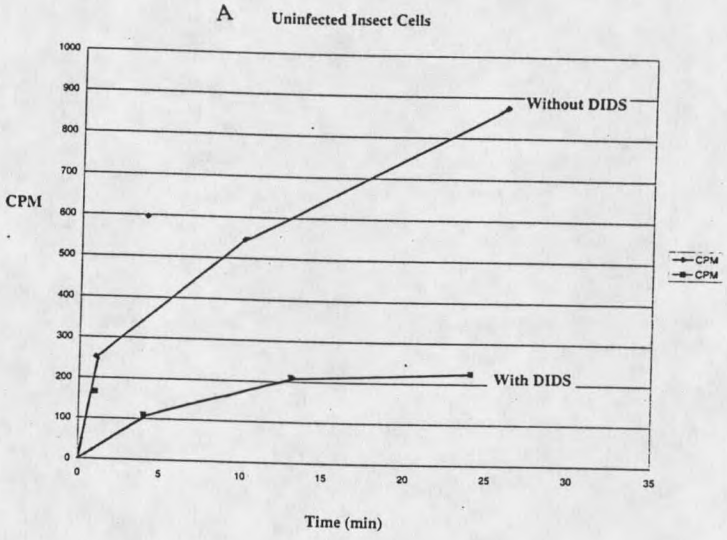

Figure 12A: **Cl-36 radioactivity assay** of chloride channels in uninfected insect cells. Note that when DIDS (a broad-spectrum chloride channel inhibitor) is added, endogenous chloride channels essentially shut down completely.

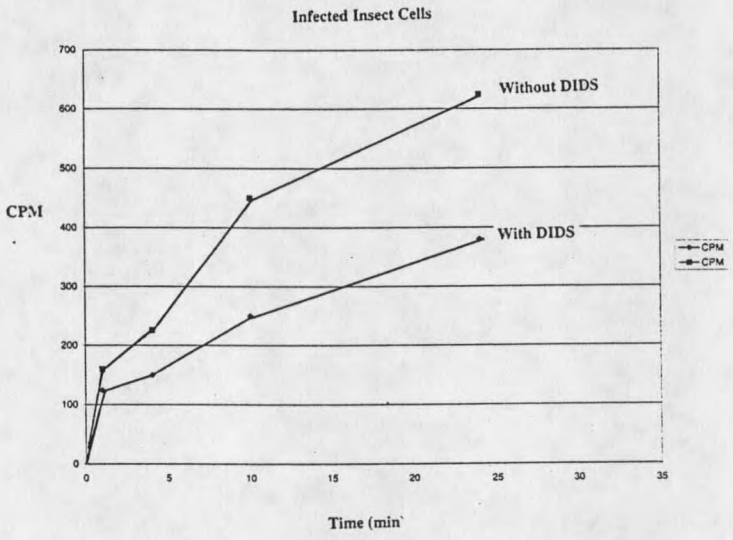

Figure 12B: **Cl-36 radioactivity assay** of chloride channels in infected insect cells expressing CFTR. Note that when DIDS is added, chloride channel activity is not entirely shut down. This is probably due to CFTR presence, which, unlike endogenous insect cell chloride channels, is not inhibited by DIDS.

above membrane vesicle fluorescence-based assay failed to detect any activity of CFTR. On a more positive note, the method described here could be useful for screening drugs for the ability to enhance CFTR activity at some point in the future.

Patch Clamp Assay of Purified CFTR

Functional reconstitution of CFTR purified from our lab was conducted in a lipid bilayer system as reported by Cantiello, et al. in 1998 (108) at Harvard (Medical School) during the Summer of 1999 in prelude to AFM experiments (Chapter 4).

Methods

Proteoliposomes containing purified CFTR were fused to planar lipid bilayers by painting them onto a 0.1 mm hole in a 13 mm polystyrene cuvette (Warner Instrument Corp) as described by Alvarez in 1986 (109). The phospholipid composition of the lipid bilayers was PE:PC:PS (7:2:1) in n-decane to a final concentration of 14 mM. The initial *cis* and *trans* solutions contained MOPS 10 mM, pH 7.4, and 10 and 100 mM MgATP, respectively (figure 13). Input signals were acquired with a PC-501A Patch/Whole Cell Clamp amplifier via a 10 Gohm head-stage for lipid bilayers (Warner Instrument Corp). The output currents were low-pass filtered at 500 Hz with an eight-pole Bessel filter, digitized at 37 kHz. Single channel tracings were Gaussian-filtered at 100-200 Hz for display purposes. Proteoliposomes were painted to the surface of the lipid bilayer and

ion channel activity was observed after addition of PKA (75 nM) to the *cis* side of the chamber.

Results

CFTR channel activity was observed in a chamber containing either 100 and 10 mM MgATP, in the *cis* and *trans* compartments, respectively, or asymmetrical Cl (cis, 200 mM) and MgATP (*trans*, 10 mM). Channel activity was also observed after addition of actin to the *cis* side of the reconstitution chamber in the absence of PKA.

Discussion

The variation on the patch clamp method described here proved capable of detecting channel activity of CFTR purified from this lab. In fact, it was the only method whereby channel activity was observed from purified CFTR in this case. Numerous attempts at purified CFTR channel activity detection using a fluorescence-based assay failed in our hands. It was eventually determined that chloride ions probably flowed much too rapidly into SPQ-loaded proteoliposomes (which are much smaller than whole cells) to detect changes in fluorescence, any of which was probably "swamped" the moment that the chloride solution was added to the cuvette. Decreases in fluorescence could not therefore be distinguished from dilution effects. As "stop-flow methods", which had the potential of correcting this situation, were not readily available to us, the patch clamp procedure described above proved to be invaluable.

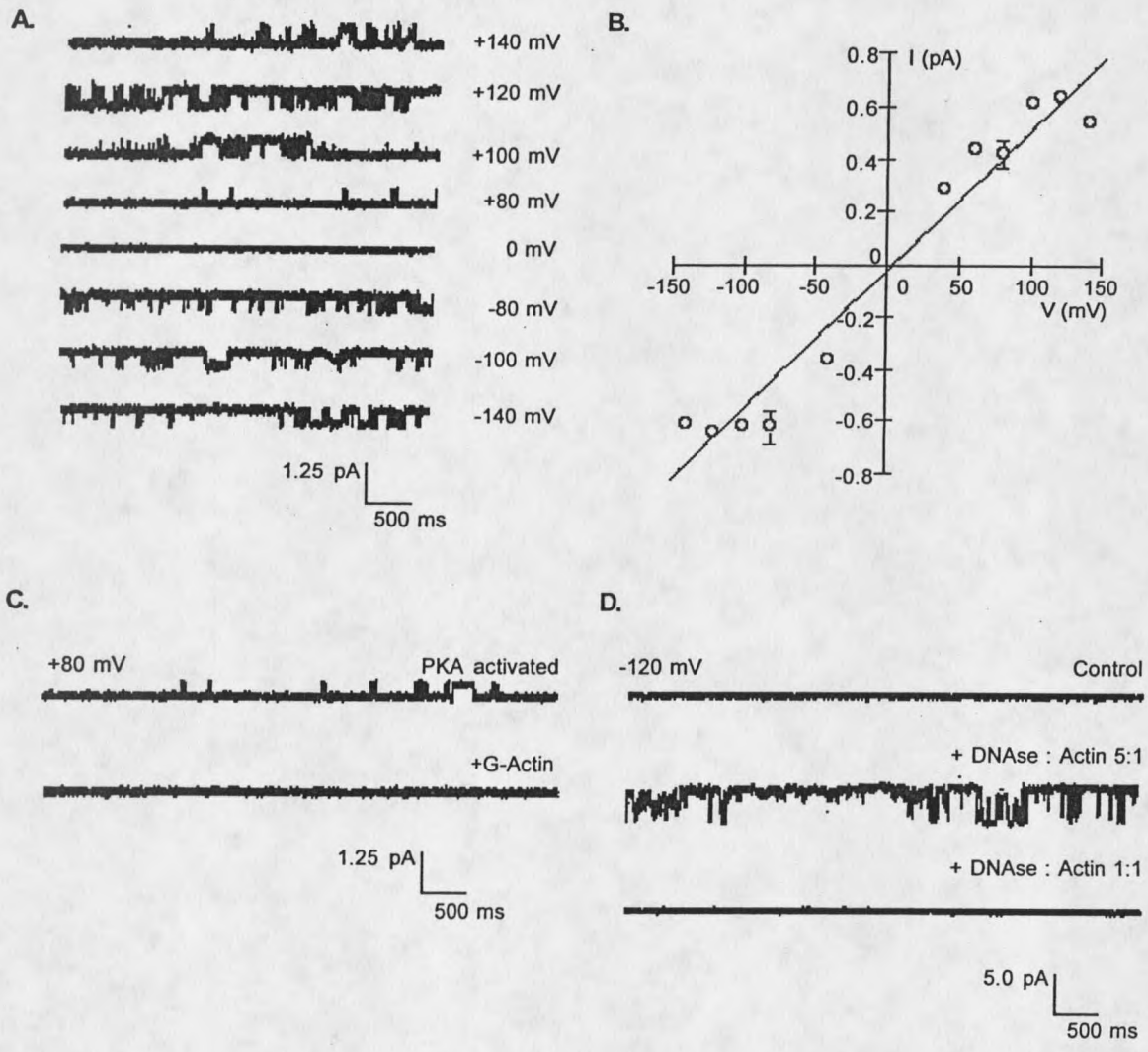

Figure 13: Patch-clamp assays of CFTR. A: single-channel recordings of CFTR from this lab, incorporated into artificial bilayers. Stochastic single-channel opening and closing events can be discerned as peaks and troughs, respectively. B: an alternative plotting of the same data in A. C: Same as in A, but with PKA (top) and with non-polymerized actin (bottom). D: DNAse is used to initiate polymerization of actin.

CHAPTER 4

CFTR-ACTIN INTERACTIONS CHARACTERIZED USING AFM

Introduction

Probably the most important aspect of ion channel study is channel gating. The CFTR chloride channel is gated (opened and closed) *in vivo* in part by phosphorylation of 6 serines at its R-domain by the kinase PKA (80). In addition, since the mid-1990s several lines of evidence have been accumulating which suggest that CFTR may be partially activated by cytoskeletal elements without the need for cellular kinase activation. For example, in 1994 Fuller et al. found that, while performing patch clamping on whole cells, CFTR activity in these colonic T84 epithelial cells was dependent to a certain extent on the structure of the microtubule cytoskeleton in the cytoplasm. Inhibitors of microtubule polymerizaton were found to be capable of slowing cyclic-AMP dependent chloride permeability (i.e. CFTR) in these cells (110). This makes sense, as microtubule disruption had been previously recognized as early as 1990 as a disrupter of the trans-Golgi network-to-apical membrane traffic in epithelial cells in (111). In 1995, Prat et al., using patch clamping, discovered that purified actin alone was capable of activating CFTR in excised membrane patches independent of kinase phosphorylation at the R-domain. Importantly, inhibitors of actin polymerization were found to disrupt this route of CFTR chloride channel activation (27). Prat et al. in 1995 found sequence similarity between an actin-binding protein and the nucleotide binding domains of CFTR (27). Again, this finding involving CFTR regulation and the cytoskeleton had been preceded by

work done by others (Cantiello et al. 1990) who had shown that epithelial Na+ ion channels are directly regulated by short actin filaments (112). Also of interest in the way of showing a putative CFTR-actin interaction, in 1997, the molecular chaperone Hsc70 was found to bind to the purified soluble NBD1 domain of CFTR by Strickland et al.(113). This may be of significance as Hsc70 is believed to be structurally related to actin (133). Still lacking as recently as 1999 was proof of a specific physical interaction between CFTR and polymerized actin filaments. It was therefore decided at that time to attempt detection of a physical actin-CFTR interaction using CFTR protein that was expressed, purified, and reconstituted from this laboratory in conjunction with atomic force microscopy (AFM) performed by Chasan et al of Boston University.

In the 15 years following the invention of AFM by Binning, Quate, and Gerber in 1986, this new surface-imaging technique has been refined sufficiently to enable the characterization of the topographic nature of several classes of biomolecules, including proteins, lipids, and nucleic acid (114). Resolution of the surfaces of biological molecules using AFM in the range of 10 Å lateral and 1 Å vertical are often attained, and forces as low as 10-50 pN arising from interactions between the tip and individual biomolecules have been measured (115).

AFM has also been used to detect physical interactions between various types of biomolecules. A significant AFM study was done by Le Cam et al. in 1994 in which TEM was used in conjunction with AFM to confirm that AFM could in fact be used to characterize a physical interaction between DNA and protein (116). The use of AFM has also enabled direct visualization of the assembly of complexes between RNA polymerase

and DNA (117). In addition, the heptameric GroES complex was visualized using AFM at 10 Å resolution before this quartenary structure was subsequently confirmed by x-ray diffraction analysis (118).

The most often used mode of operation of AFM remains the "constant force mode", whereby the cantilever tip contacts the sample continuously during a raster scan of the sample. This original method of using AFM has been found to be unsuitable for applications involving single proteins or membranes loosely adsorbed to mica surfaces (e.g. CFTR in this study), as the tip imparts excessive energy to the sample during constant force mode. Therefore, a less harsh "tapping mode" for AFM, was developed by Zhong et al. in 1993 in which, rather than the tip being in constant contact with the surface of the protein, the tip is instead oscillated on and off the sample at a frequency on the order of tens of KHz (119). Tapping mode AFM also minimizes frictional forces capable of tearing fragile proteins apart, however usually at the expense of a decrease in spatial resolution.

Some of the significant advantages that come along with the use of AFM include the ability to use physiological buffers (no vacuums are necessary), no need for stains or dyes which can lead to significant artifacts as in most EM studies, and often extremely small amounts of material (microgram or less of protein) are sufficient for detection. In these studies, nanometer-scale resolution of CFTR protein was obtained in conjunction with actin filaments using tapping mode AFM.

Methods

CFTR Expression in Insect Cells

Two flasks of T-75s were grown to near confluency with sf9 cells. After removal of media, these cells were then exposed to 5 ml baculoviral stock (MOI of 5) and allowed to incubate at room temperature 1 hour. After this time, virus was removed and replaced with fresh serum-containing media and allowed to incubate 48 hours at 27 °C. Cells were harvested by pelleting at 800 x g and frozen at –70 °C until needed. Baculovirus containing the CFTR cDNA *without* the His-tag was used for these experiments as this work preceded the tagging of the CFTR (chapter 3). This meant that it was necessary to purify CFTR using hydroxyapatite chromatography (below).

Purification of CFTR

Above pellets were thawed and insect cells lysed with alkaline extraction solution containing 3 protease inhibitors (10 µl each of leupeptin, aprotinin, and PMSF). Insect cell membranes were pelleted using ultracentrifugation. Membrane proteins were solubilized using 2% SDS detergent in phosphate buffer (10 mM, pH 7.2) and protease inhibitors. After 1 hour at room temperature, the mixture was centrifuged to remove unsolubilized membranes. Solubilized membrane proteins were added to a hydroxyapatite (ceramic type II, BioRad) column for the initial chromatography step (7). Total bed volume was 6 cm^3. Matrix was pre-equilibrated with a solution containing 10 mM phosphate, 0.5 mM DTT, and 0.15% SDS. Following the addition of protein,

unbound proteins were washed off with more of the same buffer. A step gradient was then applied (300, 320, 360, 400 mM phosphate) to remove the remainder of protein from the column. CFTR eluted from the column in the 400 mM fraction. CFTR-containing fractions were concentrated by centrifugation at 1000 x g in Centricon 50 tubes (Amicon) to a volume of about 0.5 ml. This concentrated protein was then subjected to gel filtration using G-100-120 sephadex beads (Sigma) in a column of 18 cm by 2 cm. The elution buffer contained 10 mM Tris, 100 mM NaCl, 0.25 % LiDS, pH 8.0. CFTR eluted in the first peak. This fraction was then concentrated by centrifugation (as above) to a final volume of 200 µl. This eluent was rediluted in 800 µl of a buffer containing HEPES (15 mM), EGTA (0.5 mM), pH 7.2 The diluted eluent was then re-centriconed to a final volume of 150 µl and used for reconstitution into lipid bilayers.

Reconstitution of CFTR into Artificial Lipid Bilayers

A chloroform solution containing phosphatidylethanolamine (PE), phosphatidylserine (PS), phosphatidylcholine (PC) (Avanti Polar Lipids Inc) in a 5:2:1 molar ratio was dried under Argon gas. The film was solubilized by sonication for 15 min, in a HEPES buffer (15 mM HEPES, 0.5 mM EGTA, 1% cholate) to a final lipid concentration of 10 mg/ml, pH 7.2 A 100 µl sample of sonicated lipid was then combined with the purified CFTR sample and incubated on ice for 40-60 min. This mixture was then dialyzed for 24 hours at 4 °C against two liters of HEPES buffer containing 1.5% cholate (Spectra/Por, MWCO 12-14,000), followed by daily changes of two liters of the same HEPES buffer but

without detergent for a total of 3 days. Proteoliposomes were then reconcentrated to a volume of 200 µl, frozen at –70 °C and shipped to Boston, Mass for AFM experiments.

TEM Micrographs (Figure 14)

Liposomes were formed by detergent dialysis both with (B) and without (A) CFTR present. They were double-stained (lead and uranium-based stains) and visualized after drying from pure water. Instrument used was ……..stain used…… CFTR was estimated to be completely incorporated, as unincorporated protein would have a different shape (more "straight-edged") than lipid, which is "curved-edged" using TEM.

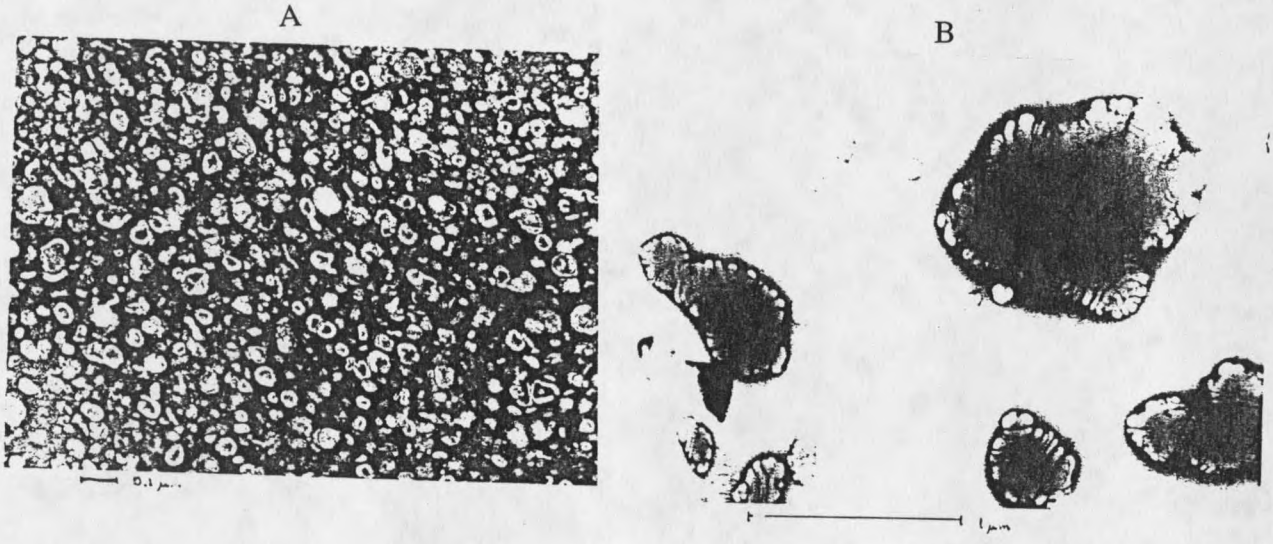

Figure 14: **TEM micrographs of liposomes** without (A) and with (B) CFTR incorporated. Note differences in size. Negative staining used was a combination of lead acetate and uranium acetate.

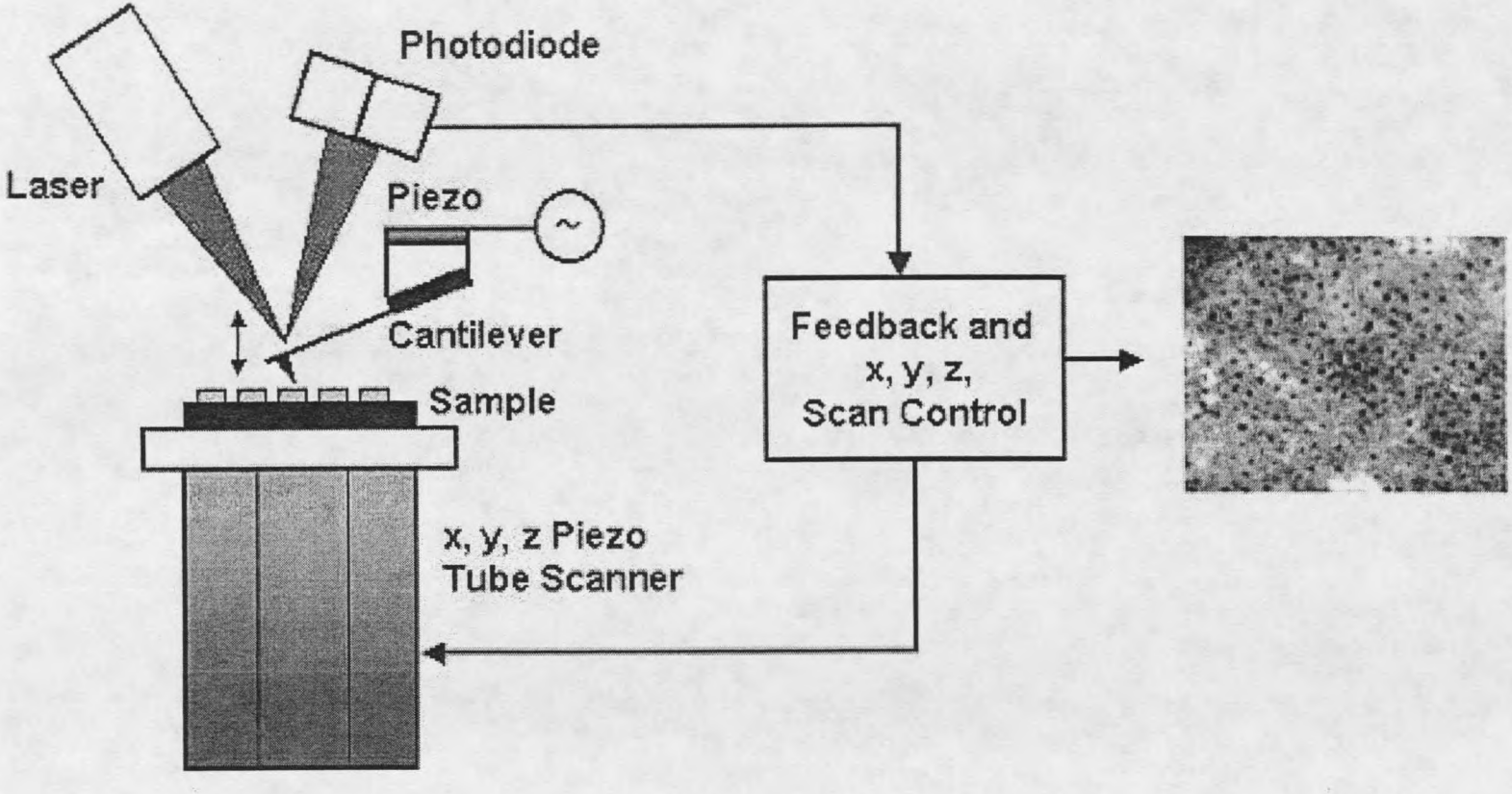

Figure 15: **Diagram of AFM.** A surface imaging technique capable of providing subnanometer scale resolution of some proteins. A fine tip is raster scanned over the sample while a laser detects changes in height.

AFM Methods

A Model 3000 AFM (Digital Instruments Inc) attached to a Nanoscope IIIa controller with an electronics extender box was used to visualize CFTR and actin (191). Tapping mode (oscillating tip), oxide-sharpened silicon-nitride tips (DNP-S; Digital Instruments) were mounted on thick 21 μm and short 100 μm cantilevers. Spring constants of cantilevers were 0.58 N/m. Frequencies of 7.9-8.2 kHz were used for tapping mode. Sampling was done on freshly cleaved mica (New York Mica Co.). All scans were in aqueous solution. Liposomes were left on mica 2 hours at 37 °C and allowed to flatten onto the mica, forming extended bilayers. Most studies were conducted at room temperature. Samples were sealed into a Petri dish kept humidified with a wet paper towel. Samples were scanned in solution by adding sufficient buffer to make a meniscus between tip and sample. 100 μl total volumes were used. Imaging was carried out in actin depolymerizing buffer containing 0.2 mM CaCl2, 0.2 mM MgATP, 0.2 mM β-mercaptoethanol and 2 mM Tris-HCl (pH 8.0). The monoclonal antibody used against CFTR was mAb 13-1 (Genzyme Corp) raised against the R domain. This mAb was diluted to a final concentration of 1:1000 for the experiments and was used to ascertain both the presence of CFTR as well as the orientation of CFTR when placed on the mica for AFM experiments (191).

Results

A diverse population of flattened liposomes with an average diameter of 82 nm (n=50) could be imaged. Liposomes fused and flattened by "baking" sample for one hour at 37 °C. The CFTR protein was found to be an oblong extramembrane protrusion with an average spherical radius of approximately 15 nm and 2 nm in height. The extramembrane molecular mass was calculated to be on the order of 320 Da. The addition of active antibody changed the size of the protruding protein and induced a 1.07 $\pm$ 0.5 nm (n = 22) increase in total height of the complex. When G-actin (i.e. non-polymerized actin) was added to the CFTR sample at a concentration of 10 nM, higher resolution scans were then done. Actin filaments were observed extending from CFTR. Little or no actin polymerization was observed detached from CFTR molecules.

Conclusions

The present study appears to indicate that physical interactions between actin and CFTR do exist under the conditions of the experiment. This supports the notion that actin can be made to bind CFTR directly and it can therefore be theorized that polymerized actin may be able to change the conformation of CFTR in such a way as to influence chloride channel gating. The calculated size of CFTR from this study was 320 kDa. As monomeric CFTR has a calculated mass of approximately 170 kDa, the data suggests CFTR is most likely forming a dimer within the membrane of the liposomes. This would support previous evidence using TEM that other eukaryotic ABC proteins such as the multidrug resistance protein Pgp are able to form dimers with one another

(120). To our knowledge, this is the first study in which a protein-protein interaction involving a transmembrane protein has been imaged by AFM (191).

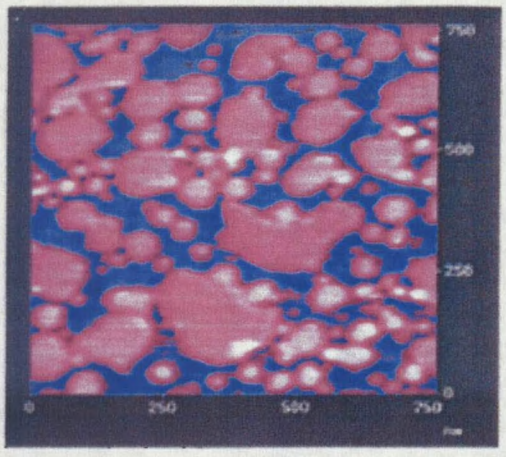

Figure 16: Liposomes (red) before baking

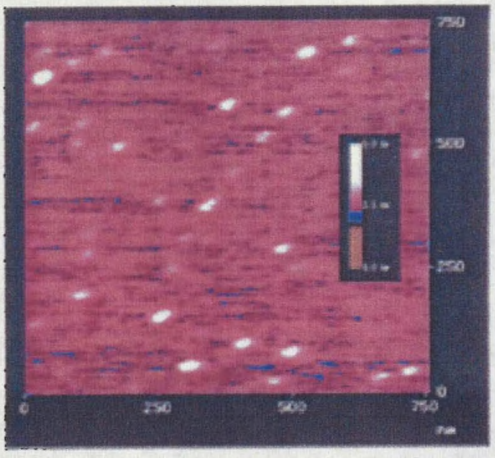

Figure 17: After baking, CFTR can be seen as white spots.

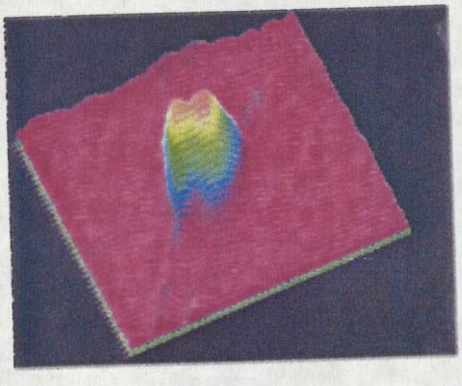

Figure 18: AFM raster scan of a CFTR dimer.

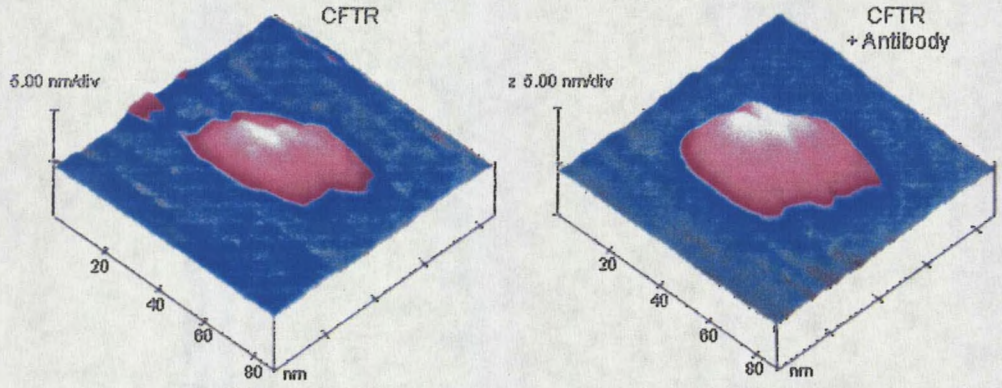

Figure 19: AFM of CFTR without (left) and with (right) addition of antibody directed against the R-domain

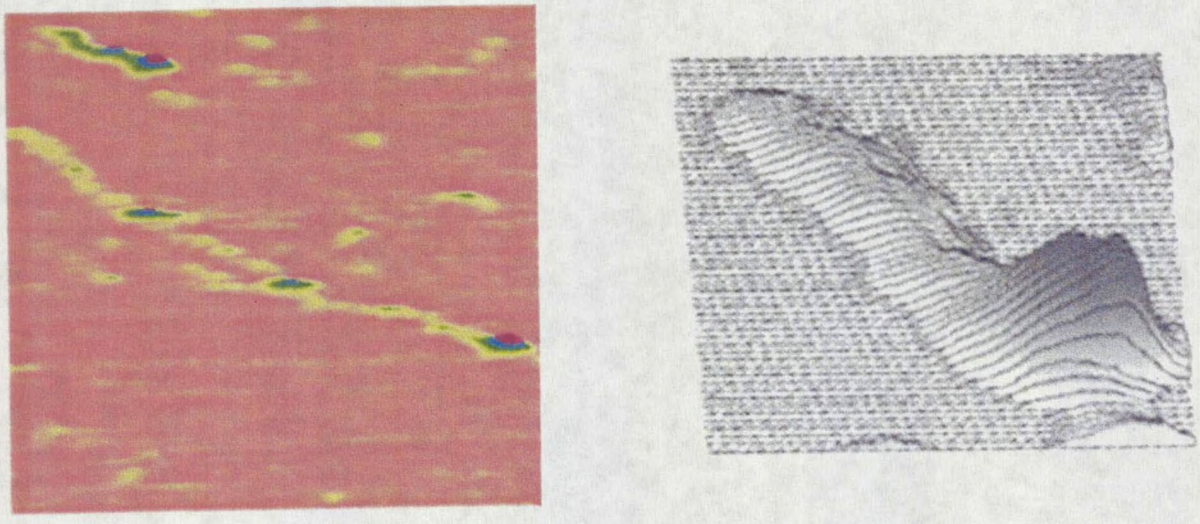

Figure 20: CFTR (green/purple) are seen connected to actin filaments (yellow). No actin filaments were found unconnected to CFTR.

CHAPTER 5

PROBING THE SOLVENT ACCESSIBILITY OF FOUR CFTR CYSTEINE RESIDUES USING CHEMICAL LABELING AND PROTEOLYSIS FOLLOWED BY MALDI-TOF MS

Introduction

Cystic fibrosis is the most common lethal genetic disease in North America, and patients have a mean survival age of only 30 years. The gene responsible for cystic fibrosis is known to code for a chloride channel and regulator called the cystic fibrosis transmembrane conductance regulator, or CFTR. It has become increasingly clear during the last decade since its discovery that a there exists a need for additional methods capable of characterizing the structure of CFTR and other transmembrane proteins. Most of the structural data gathered concerning CFTR over the last 10 years has come about as a result of such low-resolution techniques as CD spectroscopy (127-131), native polyacrylamide gel electrophoresis (132-133), sucrose density gradient centrifugation (134), Fourier Transform Infrared Spectroscopy (130), and mutagenesis (135-139). In addition, the majority of these studies have necessitated the use of prokaryotically-produced, water-soluble polypeptides based on the predicted domains of CFTR (NBD1, NBD2, RD), or short synthetic peptides based on the CFTR transmembrane helices (140-141) (130), as opposed to the intact, native CFTR protein molecule. Studies using higher-resolution techniques such as 15N, 1H-NMR are reportedly ongoing using a bacterially-produced CFTR domain (NBD1), and NMR has been applied to peptides synthesized based on portions of the NBD1 sequence of CFTR, however no high-

resolution structure currently exists for NBD1 (142) (140) or for any other portion of CFTR. In short, very little information has been gained concerning the 3-dimenstional structure of the ion channel beyond the topological orientation first predicted for it in 1989, which was based solely on the primary sequence.

In an effort to bring mass spectrometry-based methods to bear on whole native CFTR protein, a method involving the use of MALDI-TOF MS was used here for the identification of 80 proteolyzed CFTR peptides extending over 41% of the primary sequence of CFTR. After 9 separate digests using Staphylococcal V8 protease and trypsin, peptides from all 5 domains of CFTR were identified, and isotopic resolution was obtained in reflectron mode for all but 4 of these peptides. In an effort to gain more structural information using this mass spectrometry-based method, chemical labeling was also used to modify solvent-accessible Cys residues.

It was found that only 1 of the 4 Cys-containing peptides recovered after labeling in the native state, followed by denaturation and proteolysis, had been alkylated with the hydrophilic reagent iodoacetic acid (IAA), suggesting this portion of CFTR is exposed to the solvent. Interestingly, the IAA-labeled residue, Cys-343, has been previously proposed to be located within the TM6 portion of CFTR. This may be due to Cys-343 being on the face of the TM6 helix that lines an aqueous channel. Three additional cysteine residues (Cys-524, Cys-647, and Cys-1410) were not found to be accessible to the solvent because they did not react with the hydrophilic reagent IAA. These residues were located within the predicted the NBD1 domain, the R-domain and the C-terminal tail region of CFTR respectively.

The method reported here was performed without the need for chromatographic separation of the peptides after digestion. It also did not require extensive sample preparation prior to mass spectrometry (i.e. buffer and detergent removal). This is significant because sample preparation of hydrophobic membrane protein digests for mass spectrometry has been shown to reduce the overall yield of peptides by up to 1000-fold (144).

MALDI-TOF MS has been used successfully for several aspects of protein structure elucidation, most recently in the identification of specific protein folds (145), the identification of disulfide bond placement within proteins (146), glycosylation site identification (147-148), phosphorylation site characterization (149), the location of protein-protein interactions (protein foot-printing) (150-151), the monitoring of changes in protein structure due to allosterism (152), as well as epitope mapping of monoclonal antibodies (153). Recently, as little as 30 pmol of the transmembrane protein rhodopsin was analyzed using a similar method on the same Bruker Biflex MALDI instrument used here, with the resulting CNBr-produced peptides giving coverage of 99% of rhodopsin' s primary sequence (154).

Methods

Materials

Sf9 cells, TNM-FH media, and baculovirus (AcNPV) were from Invitrogen (Carlsbad, CA). Bovine pancreatic trypsin type III and V8 protease type XVII-B were

from Sigma (St. Louis, MO)., 7.5 % SDS-PAGE gel and Zn stain were from BioRad; (Hercules, CA), Protease inhibitors: aprotinin, leupeptin, and PMSF were from Sigma (St. Louis, MO); lysophosphadidyl glycerol (LPG) was from Avanti Lipids (Alabaster, AL). α-cyano-4-hydroxycinnamic acid matrix (ACHA) was from Aldrich Chem. Co. (Milwakee, WI), and NANOSEP MF microcentricon tubes from PALL FILTRON (Northborough, MA). The MALDI-TOF MS instrument was a Bruker Biflex III Protein Prospector (UCSF). HisBind Resin™ used for CFTR purification (Novagen; Madison, WI). IAA (Sigma; St. Louis, MO)

Expression and Purification of CFTR.

Sf9 cells were grown to confluency in 15 T-75 flasks and infected while in log phase with baculovirus incorporating CFTR cDNA expressing a polyHIS tag to aid in purification (see chapter 3). Cells were harvested 48 hours after infection. Cell pellets were then lysed with alkaline phosphatase buffer (note: this procedure was modified when treating CFTR first with IAA, see below) containing the following protease inhibitors: (leupeptin, aprotinin, PMSF, 50 μg/ml each final concentration). Cell membranes were then solubilized with 2 ml of Buffer A (10 mM Tris, pH 7.9) containing 2% LPG and protease inhibitors. Nickel chromatography is a type of affinity chromatography which works according to the following theory. First, a protein of interest is designed to have incorporated onto its tail several consecutive histidines. This protein can be purified away from other proteins not containing these residues by simply allowing it to bind to the ionic form of ionic nickel. This transition metal has

been attached to a matrix such as sepharose. In this case, the procedure was performed as a slurry initially. Briefly, 1-3 ml (settled volume) of suspended beads were charged with nickel sulfate as per instructions in the Novagen pET System manual (7th Ed.). Detergent-solubilized membrane protein was added to a slurry of nickel-charged beads and rotated in a slow tumbling motion for 5 hours at 4 °C. After 5 hours, the slurry was transferred to a chromatography column and beads were allowed to settle. The first elution step was performed with 2 column volumes of binding buffer (5 mM imidazole in Buffer B: 15 mM Tris, 500 mM NaCl, 0.02% LPG, pH 7.9). After rinsing, this column washing was repeated for 3 more steps (60 mM, 430 mM, and 1000 mM imidazole all in buffer B) 2 column volumes each. All 3 samples were concentrated to a volume of approximately 400 µl using ultracentrifugation and tested for presence of CFTR using SDS-PAGE/silver stain and western blot (results not shown) (chapter 3).

Cell Membranes Treated With IAA:

The above procedure was performed when treating cells with IAA with the following exceptions: CFTR-expressing cells were lysed using a 27 G needle after harvesting cells instead of with alkaline extraction solution. The lysis buffer used consisted of TBS (135 mM NaCl, 2.5 mM KCl, 25 mM Tris, pH 8.6) and 200 mM IAA. Volume proportions were approximately 5 ml wet cell pellet to 25 ml IAA-containing buffer. After lysing Sf9 cells with the needle in the dark, membranes were pelleted using an ultracentrifuge (35K, 40 min). This step was done in just under 20 minutes in an effort to minimize contact of CFTR with soluble proteases. The cell membrane pellet was then

resuspended in 10 ml additional buffer containing IAA and allowed to react on ice in the dark for 2 hours to allow all available Cys residues to react with IAA. IAA was added again because the initial exposure to IAA lasted only about 20 minutes. Cell membranes were then spun again as above to wash off any remaining unreacted IAA and resuspended in alkaline extraction buffer for approximately 5 minutes for the removal of peripheral membrane proteins. Cell membranes were then ultracentrifuged again and the resulting membrane pellet rinsed on the surface briefly with deionized water.

Gel Preparation for MALDI:

A standard SDS-PAGE gel (7.5% polyacrylamide, Tris/SDS buffer, with a stacking layer of 4% acrylamide) was loaded in all available lanes with 50 µl/lane of ~ pure CFTR (total CFTR in gel was estimated to be up to 600 pmol, or between 50 and 100 µg total). Gel was run at 150 V until dye front neared bottom of gel. Gel was then rinsed briefly in deionized water and then subjected to zinc staining. In this procedure, the polyacrylamide gel is subjected to a solution containing histidine for about 10 minutes. This amino acid diffused throughout the gel except where protein is present. Next, the gel is exposed to a clear solution containing zinc in the ionic form. The zinc next binds to the histidine in the buffer and turns opaque. After 10 minutes in the zinc-containing solution, the gel is removed and placed against a black background. This allows the clear protein bands to be visualized more easily. Following this procedure, a continuous protein band across the width of the gel approximately 8 mm in width was observed following at the apparent molecular weight CFTR migrates at (135 KD). This

negatively-stained CFTR band was then removed with a razor blade and destained of any excess zinc. This destaining step was performed twice.

CFTR band preparation for MALDI:

The CFTR-containing gel strip was frozen at –70 °C and then lyophilized to dryness (2 hours) and then rehydrated in 10% acetonitrile/water. After saturating the gel strip, an additional 1 ml acetonitrile solution was added in excess. The gel strip was next tumbled for 15 minutes at room temperature. Afterwards, the acetonitrile solution (containing residual SDS) was removed and the gel strip was again frozen at –70 °C and then lyophilized to dryness (4 hours). For digestion of CFTR, either bovine pancreatic trypsin (3.5 mg/ml solubilized first in deionized water) or V8 protease was solubilized to a final concentration of 0.1 mg/ml in ~5 mM Tris pH 8.0. For trypsin preparation, this stock solution was then diluted in 10% acetonitrile/water to give a final trypsin concentration of 0.1 mg/ml. The dried and lyophilized gel strip was then allowed to take up approximately 200 µl of either the V8 or trypsin protease-containing solution 50 µl at a time by capillary action until the gel appeared saturated. Afterwards, 50 µl excess of the solution was added and tube was incubated at 37 °C overnight in a water bath. To remove the peptides from the gel for MALDI analysis, the gel slice was next placed in a NANOSEP MF microcentricon tube. The gel strip was then centrifuged on high speed at room temperature approximately 15 minutes to recover peptides from the gel. Lastly, 1 µl of this supernatant was added to 1 µl matrix solution (saturated ACHA in 70% acetonitrile) and half a microliter of this mixture was spotted onto a MALDI sample plate.

Additional CFTR peptides were recovered when the above procedure was modified in the following way. A 5x5 mm fragment of the gel spun in the NANOSEP MF microcentricon tube (above) was cut out of the gel strip with a spatula and soaked in approximately 10 µl of 100% acetonitrile. After several minutes, excess acetonitrile solution with the diffused CFTR peptides was subjected to MALDI as described above.

MALDI-TOF:

All spectra were obtained on a Bruker BiFlex III. Typical accelerating voltages were

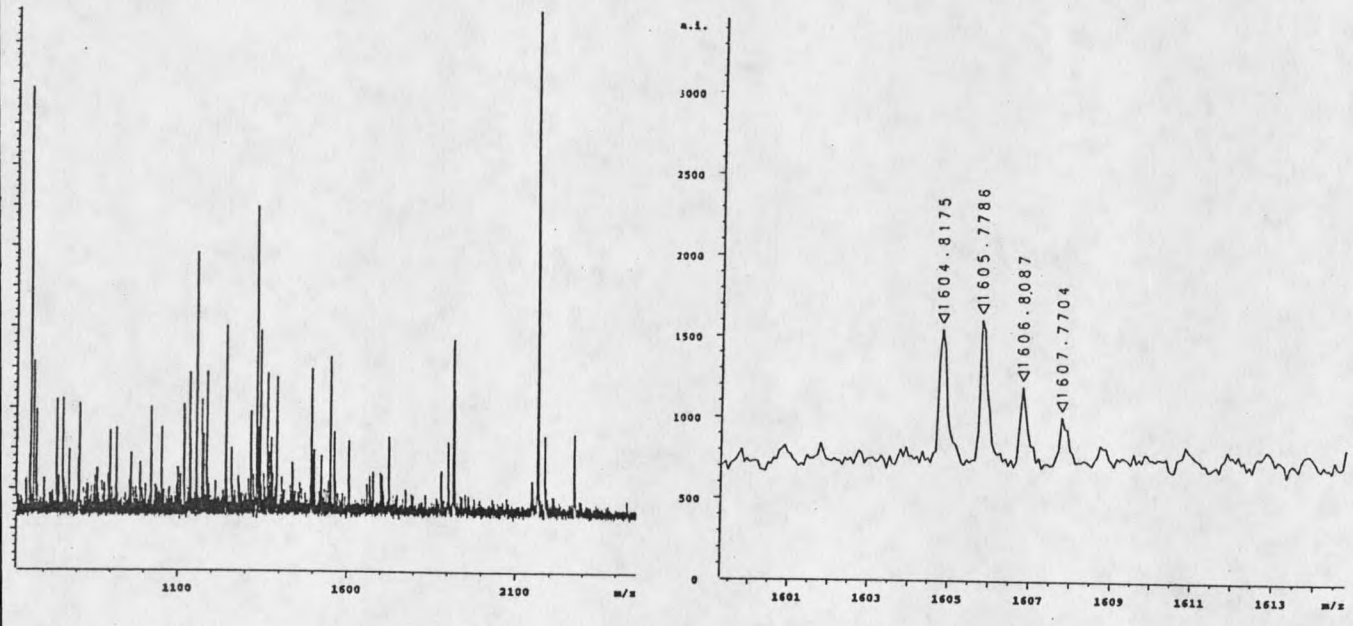

Figure 21: (left) **MALDI-TOF MS spectrum** of trypsin in situ digest of whole CFTR in reflectron mode. (right) ENIIF*GVSYDEYR peptide (blow-up of peak from spectrum on left). Peptide is from NBD1. Shows isotopic resolution. F* indicates Phe-508, a critical amino acid deleted in almost 90% of CF patients.

20 kV, and delayed extraction times ranged from 100 to 400 ns. Spectra were averaged over 100 to 600 laser shots. A linear calibration was applied to all spectra using trypsin or V8 autoproteolysis fragments and matrix dimer (379 amu) as an internal calibration points to maximize mass accuracy.

Results

MALDI-TOF MS spectra were all performed on peptides extracted from CFTR digests after addition of matrix. Peptide masses were matched to those in the theoretical digests using the program MS Digest (Protein Prospector, http://prospector.ucsf.edu). Following mapping of peptides onto the primary sequence of CFTR, it was noticed that the majority of the peptides from the *in gel* CFTR digests were located in the primary sequence either adjacent to or partially overlapping other peptides (figure 24). In addition, some regions of CFTR appeared to be more susceptible to proteolysis than others, such as the two NBD domains, the R-domain and intracellular loop 2 (ICL2), however it was not possible to rule out other causes such as the ionization process during MALDI causing certain regions to be over-represented in the spectra.

The majority of peaks in most of the digests were readily assignable to either CFTR or protease autoproteolysis products (often used for calibration). Of the 76 peptides from 9 digests assigned to CFTR in reflectron mode, 66 had errors in the peptide masses of less than 60 ppm, as shown in tables 1 & 2. The remaining 10 masses with errors greater than 60 ppm are believed to be peptides from CFTR for two reasons. First, these