(19) [Canadian flag symbol] **Canadian Intellectual Property Office**

An Agency of Industry Canada

Office de la Propri,t, Intellectuelle du Canada

Un organisme d'Industrie Canada

(11) **CA 2 091 907**   (13) **C**

(40) **04.02.1992**
(43) **04.02.1992**
(45)

(12)

(21) **2 091 907**

(22) **20.09.1991**

(51) Int. Cl.^{5}: **C12N 15/12, C12N 5/10, C12N 1/21, C07K 14/47**

(85) **17.03.1993**

(86) **PCT/CA91/00341**

(87) **WO92/05252**

(30)   **90 20632.7   GB   21.09.1990**

(72)
      TSUI, Lap-Chee (CA).

      ROMMENS, Johanna M. (CA).

(74)
      SIM & MCBURNEY

(54)   PROPAGATION STABLE DE L'ADN COMPLEMENTAIRE MODIFIE SUR TOUTE SA LONGUEUR D'UNE PROTEINE DE REGULATION DE LA CONDUCTANCE MEMBRANAIRE DE LA FIBROSE KYSTIQUE DANS DES SYSTEMES HETEROLOGUES

(54)   STABLE PROPAGATION OF MODIFIED FULL LENGTH CYSTIC FIBROSIS TRANSMEMBRANE CONDUCTANCE REGULATOR PROTEIN CDNA IN HETEROLOGOUS SYSTEMS

(57)
        A modified DNA sequence encoding full length cystic fibrosis transmembrane conductance regulator protein is provided to facilitate propagation and/or expression of the protein in living cells and in particular, bacterial cells. The modified DNA sequence comprises at least one of the 13 base pair repeat of exon 6b of the normal gene encoding the conductance regulator pro- tein, as one or more normal nucleotides of the 13 base pair repeat substituted with an alternate nucleotide which however conti- nues to code for the corresponding normal amino acid. Mammalian cells transfected with a vector containing the modified DNA sequence enhances chlorine conductance through the cell wall.

(54) Titre : PROPAGATION STABLE DE L'ADN COMPLEMENTAIRE MODIFIE SUR TOUTE SA LONGUEUR D'UNE PROTEINE DE REGULATION DE LA CONDUCTANCE MEMBRANAIRE DE LA FIBROSE KYSTIQUE DANS DES SYSTEMES HETEROLOGUES
(54) Title: STABLE PROPAGATION OF MODIFIED FULL LENGTH CYSTIC FIBROSIS TRANSMEMBRANE CONDUCTANCE REGULATOR PROTEIN cDNA IN HETEROLOGOUS SYSTEMS

(57) Abrégé/Abstract:
A modified DNA sequence encoding full length cystic fibrosis transmembrane conductance regulator protein is provided to facilitate propagation and/or expression of the protein in living cells and in particular, bacterial cells. The modified DNA sequence comprises at least one of the 13 base pair repeat of exon 6b of the normal gene encoding the conductance regulator protein, as one or more normal nucleotides of the 13 base pair repeat substituted with an alternate nucleotide which however continues to code for the corresponding normal amino acid. Mammalian cells transfected with a vector containing the modified DNA sequence enhances chlorine conductance through the cell wall.
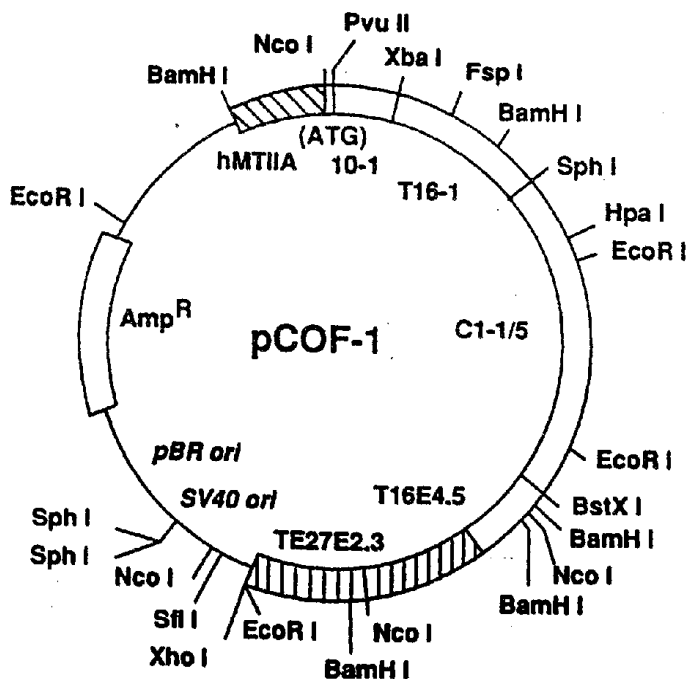
(51) International Patent Classification 5 :		(11) International Publication Number: WO 92/05252
C12N 15/12, 1/21, 5/16 **2091907** C07K 13/00	A1	(43) International Publication Date: 2 April 1992 (02.04.92)

(21) International Application Number: PCT/CA91/00341

(22) International Filing Date: 20 September 1991 (20.09.91)

(30) Priority data:
9020632.7   21 September 1990 (21.09.90) GB

(71) Applicant (for all designated States except US): HSC RESEARCH AND DEVELOPMENT LIMITED PARTNERSHIP [CA/CA]; 88 Elm Street, Toronto, Ontario M5G 1X8 (CA).

(72) Inventors; and
(75) Inventors/Applicants (for US only) : TSUI, Lap-Chee [CA/CA]; 94 Willowbridge Road, Toronto, Ontario M9R 3Z4 (CA). ROMMENS, Johanna, M. [CA/CA]; 199 Bogert Avenue, Willowdale, Ontario M2N 1L1 (CA).

(74) Agent: WOODLEY, John, H.; Sim & McBurney, 330 University Avenue, Toronto, Ontario M5G 1R7 (CA).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GB (European patent), GR (European patent), IT (European patent), LU (European patent), NL (European patent), NO, SE (European patent), US.

Published
With international search report.

(54) Title: STABLE PROPAGATION OF MODIFIED FULL LENGTH CYSTIC FIBROSIS TRANSMEMBRANE CONDUCTANCE REGULATOR PROTEIN cDNA IN HETEROLOGOUS SYSTEMS

(57) Abstract

A modified DNA sequence encoding full length cystic fibrosis transmembrane conductance regulator protein is provided to facilitate propagation and/or expression of the protein in living cells and in particular, bacterial cells. The modified DNA sequence comprises at least one of the 13 base pair repeat of exon 6b of the normal gene encoding the conductance regulator protein, as one or more normal nucleotides of the 13 base pair repeat substituted with an alternate nucleotide which however continues to code for the corresponding normal amino acid. Mammalian cells transfected with a vector containing the modified DNA sequence enhances chlorine conductance through the cell wall.

STABLE PROPAGATION OF MODIFIED FULL LENGTH CYSTIC FIBROSIS TRANSMEMBRANE CONDUCTANCE REGULATOR PROTEIN cDNA IN HETEROLOGOUS SYSTEMS

FIELD OF THE INVENTION

This invention relates to modifications in the cDNA of full length cystic fibrosis transmembrane conductance regulator protein (CFTR) which facilitates propagation and/or expression in heterologous systems.

5   ## BACKGROUND OF THE INVENTION

Cystic fibrosis (CF) is the most common, life-threatening, autosomal recessive disease in the Caucasian population. Approximately 1 in 2,500 live births is affected by this genetic disorder. Obstructive lung disease, pancreatic enzyme insufficiency and elevated sweat electrolytes are the hallmarks for CF

10   but the severity of these symptoms vary from patient to patient. Patients with CF usually die at an early age due to lung infection. With recent advances in clinical treatments, which are directed against the symptoms, the mean survival age for patients has increased to 26 years.

Despite intensive research efforts for the past fifty years, the basic

15   defect in CF remains to be speculative. It is generally believed that the heavy mucus found in the respiratory tracts and the blockage of exocrine secretion from the pancreas are due to imbalance in water secretion which is the consequence of a defect in the regulation of ion transport in the epithelial cells.

20   The precise localization of the CF locus on the long arm of chromosome 7, region q31, facilitated the recent isolation of the responsible gene. The CF gene spans 250 kilobase pairs (kb) of DNA and encode a mRNA of about 6,500 nucleotides in length. The CFTR gene is disclosed in United States Patent No. 5,981,178.

2

Expression of this gene could be observed in a variety of tissues that are affected in CF patients, for example, lung, pancreas, liver, sweat gland and nasal epithelia. An open reading frame spanning 1480 amino acids could be deduced from the overlapping cDNA clones isolated. The putative protein

5   as noted is called "Cystic Fibrosis Transmembrane Conductance Regulator" or CFTR for short, to reflect its possible role in the cells. The predicted molecular mass of CFTR is about 170,000.

Based on sequence alignment with other proteins of known functions, CFTR is thought to be a membrane-spanning protein which can function as a

10   cyclic AMP- regulated chloride channel. The internal sequence identity between the first and second half of CFTR resembles the other prokaryotic and eukaryotic transport proteins, most notably, the mammalian P-glycoprotein.

The most frequent mutant allele of the CF gene involves a three base

15   pair (bp) deletion which results in the deletion of a single amino acid residue (phenylalanine) at position 508, within the first ATP-binding domain of the predicted polypeptide. Although this mutation (ΔF508) accounts for about 70% of all CF chromosomes, there is marked difference in its proportion among different populations. The remaining 30% of mutations in the CF gene

20   appear to be heterogeneous and most of them are rare, with some represented by only single examples.

The mutation screening study confirms that the ATP-binding domains detected by sequence alignment is important for CFTR function as multiple, different mutations have been found for many of the highly conserved amino

25   acid residues in these regions. The locations of the various mutations also identified other functionally important regions in CFTR. There is, for example, a section three bp deletion resulting in the

omission of an isoleucine residue at position 506 or 507 of the putative protein. While amino acid substitutions at these positions are apparently not disease-causing, this observation argues that the length of the peptide is

5 more critical than the actual amino acid residue in the 506-508 region. Further, the existence of a large number of nonsense, frameshift as well as mRNA splicing mutations in the CF gene implies that absence of CFTR is not incompatible with life.

10 The varied symptoms among different CF patients suggest that disease severity is at least in part related to the mutations in the CF gene. Such association, which is expected to be concordant among patients within the same family, as they should have the same genotype at the

15 CF locus, is observed for pancreatic function. Approximately 85% of CF patients are severely deficient in pancreatic enzyme secretion, thus diagnosed as pancreatic insufficient (PI), and the other 15% have sufficient enzyme, thus pancreatic sufficient (PS).

20 Family studies showed that there was almost complete concordance of the pancreatic status among patients within the same family, leading to the suggestion that PI and PS are predisposed by the patients' genotypes. Subsequent studies showed that patients homozygous for

25 the ΔF508 mutation were almost exclusively PI. This information may be useful in disease prognosis.

There are other mutations that would be classified in the same group as ΔF508, the so-called severe mutant alleles with respect to pancreatic function. In

30 contrast, patients with one or two copies of other class (i.e. mild) of alleles are expected to be PS. Meconium ileus which is observed in about 30% of CF patients appears to be a clinical variation of PI and not directly determined by the CF genotype. Other clinical

35 manifestations are more complicated and no apparent association has yet been detected.

2091907₄

With the identification of the CF gene, a better understanding of the basic defect and pathophysiology of the disease can now be attained. Progress and advance are being made in studies of the regulatory mechanisms

5 governing the expression of this gene, and of the biosynthesis and subcellular localization of the protein (through generation of antibodies against various parts of the protein). In addition, it is important to develop effective assay systems for the function of CFTR. This

10 information may be useful in development of rational therapies, including gene therapy.

In order to obtain a DNA sequence containing the entire coding region of CFTR, it is necessary to construct a full-length cDNA from overlapping clones

15 previously isolated. A major difficulty has been encountered in the process, however. As the various proportions of the full-length cDNA is being linked together by standard procedures; i.e., restriction enzyme cutting and ligation, with plasmid vector in Escherichia

20 coli, frequent sequence rearrangement has been detected in the resulting construct.

For purposes of better understanding of the regulatory functions of the CFTR protein and also for purposes of gene and drug therapy, it is useful to be

25 able, in a commercial way, to propagate and express the normal CFTR gene and various mutant CFTR genes in a variety of hosts which include bacteria, yeast, molds, plant and animal cells and the like.

Although propagation and expression of the cDNA

30 sequence for the CFTR gene can be achieved in some vehicles, there are, however, the aforementioned difficulties in obtaining stable propagation of the cDNA in some types of bacteria, particularly E. coli. It is thought that the cDNA contains sequence portions which,

35 when propagated in the bacteria, results in a toxic effect which is countered by lack of propagation of the cDNA in the microorganism.

5

INVENTION SUMMARY

We have discovered that a modification of a cDNA repeat sequence in exon 6 of the CFTR gene without modifying the amino acids encoded by the changed codons facilitates propagation and/or expression of the CFTR

5   protein in living cells and in particular bacterial cells.

According to an aspect of the present invention, there is provided a modified DNA sequence encoding a cystic fibrosis transmembrane conductance regulator (CFTR) protein, said gene having at least 27 exons of normal cDNA encoding a CFTR protein of Figure 1, said normal cDNA

10   including exons 6a and 6b wherein exon 6b includes a 13 bp repeat,

said modified DNA sequence comprising at least one of said 13 bp repeats of exon 6b having one or more nucleotide substitutions in said 13 bp repeat wherein the amino acid sequence of the protein encoded by said DNA sequence is not changed by said one or more substitutions.

15   According to another aspect of the present invention, a DNA construct for use in a recombinant vector comprises the modified cDNA.

According to a another aspect of the present invention, a vector which comprises the DNA construct and a promoter sequence for the DNA construct.

20   According to further aspect of the present invention, a host cell for producing CFTR protein, the host cell comprises the above vector whereby expression of the vector in the host cell produces CFTR protein.

According to another aspect of the present invention, a CFTR protein isolated and purified from culture of the host cell.

25   According to a further aspect of the present invention, a mammalian cell transfected with the above vector to enhance C1 conductance through a cell wall.

DESCRIPTION OF DRAWINGS

Figure 1 is the nucleotide sequence of the CF gene and the amino acid sequence of the CFTR protein;

Figure 2 is a schematic diagram of the CF gene and
5   its protein product.   (A) Gene structure with exons represented by open boxes; (B) Computer-predicted primary structure of CFTR; and (C) Model of CFTR showing possible relationship with cellular membrane and potential glycosylation site (-CHO);

10   Figure 3 is a DNA sequence of exon 6b.   The nucleotide positions (877-1002) for exon 6 correspond to the previous cDNA numbering scheme (Riordan et al 1989). The 13 bp direct repeats are underlined.   The consensus prokaryotic transcription signals (at positions -35 and -

15   10) are included for comparison.   The modified nucleotides in pCOF-1; pBQ6.2 and pBQ4.7 are shown underneath.

Figure 4 is a schematic diagram showing the three basic steps in the construction of full-length CFTR cDNA.

20   Figure 5 is the generation of the 5' end of CFTR cDNA for the expression vector pCOF-1.   Two overlapping, oppositely oriented oligonucleotides, Cof1 (44-mer) and Cof2 (40-mer), were used in the generation of the exon 1 sequence.   A double stranded molecule was synthesized

25   from these two oligonucleotides with the Klenow fragment of DNA polymerase.   An NcoI site was introduced at the initiation codon to facilitate subsequent cloning.   The PCR-generated fragment was digested with NcoI and PvuII and ligated to the remaining portions of the CFTR cDNA,

30   including the PvuII to XbaI fragment of cDNA clone 10-1, the PCR-modified middle fragment and the 3' fragment . The procedures used to generate the two latter segments are described with reference to Figures 6 and 7.   The position of the human metallothionein IIA promoter, the

35   bacterial ampicillin-resistance gene, the pBR origin of replication and the SV40 origin of replication are marked.

7

Figure 6 is 3'end construction. The 3' coding region of the expression vector (pCOF-1) and pBQ6.2 was constructed in two sequential cloning steps. First, the 3' most portion, including the end of the coding sequence and the 3' untranslated region, was prepared by ligating a fragment from T16-4.5

5   (generated by BcoI partial digestion) to a fragment from the genomic clone TE27 (Rommens et al 1989). The resulting fragment was then joined with the SphI—BstXI fragment from C1-1/5 (Riordan et al 1989) at the BsXI restriction site to yield a 4 kb 3' end fragment.

Figure 7 is a schematic diagram showing the synthesis of the middle

10   segment containing the modified sequence. Details are described herein.

Figure 8 is propagation vectors pBQ4.7 and pBQ6.2. The construction of these plasmids are provided in description of the invention. Key restriction sites are marked. The sizes of the transcripts generated by the T7 RNA polymerase (from the PstI site to each of the marked restriction sites) are

15   also indicated (in kilobases).

Figure 9 is DNA sequencing ladder showing the modified sequence in exon 6b. The sequencing reaction was performed with the use of an oligonucleotide primer (19-mer) corresponding to a sequence in exon 6a starting at nucleotide 771 (5'-GCTAATCTGGGAGTTGTT-3'). The altered

20   nucleotide in the sequence are underlined.

Figure 10 is in vitro translation of CFTR. The products from in vitro translation reactions were separated on a 10% polyacrylamide -SDS gel. The gel was stained by Coomassie Blue after electrophoresis; it was then soaked in Enhance® (New England Nuclear) and dried at 60°C under vacuum before

25   exposure to X-Omat™ film (Kodak). Lane 1: no added RNA; lane 2: protein translated from the Brome Mosaic Virus RBA (as size and reaction conditions control); lanes 3 and 4: duplicate samples of capped mRNA prepared by T7 RNA polymerase on template pQB6.2 digested with HpaI. The positions for

8

the protein molecular weight standards (purchased from BRL Labs) are indicated on the left in kilodaltons. The expected product of 82.6 kb (from the RNA generated from the pBQ6.2 plasmid) is marked. Also marked (by arrows) are the sizes of the BMC protein products.

5   Figure 11 is DNA hybridization analysis of integrated plasmids. Genomic DNA samples were prepared from each cell line, digested with BamHI and separated on a 0.7% agarose gel in Tris-acetate buffer. The DNA lanes were transferred to Zetabind™ (Bio-Rad Labs) membrane. The autoradiograph shows the hybridization result with the CFTR cDNA probe.

10   Lanes 1 and 2 contain human and mouse genomic DNA, respectively. Lanes 3, 4, 25 and 26 contain non-transfected mouse LTK-cell DNA. Lanes 5 to 9 contain DNA from HAT-resistance clones obtained by transfection with pSTK7 only. Lanes 10 to 18 contain DNA from HAT resistant clones obtained by co-transfection with pSTK7 and pCOF-1 and lanes 19 to 24 with pSTK7 and

15   pCONZ. Molecular weight markers are shown on the left (in kb). The two diagnostic bands are 2.3 kb (the 5' end of the CFTR gene plus the expression vector, promoter sequence) and 2.5 kb (CFTR internal segment); they are indicated on the right by the arrows.

     Figure 12 is RNA hybridization analysis. Total RNA was extracted from

20   each cell line, purified through CsC1 gradient centrifugation, and separated on a 1.0% formaldehyde agarose gel in MOPS-acetate buffer. The samples were transferred to Zetabind (Bio-Rad Labs) membrane and hybridized sequentially with radioactively labeled CFTR cDNA (panel A) and the TK cDNA probe (panel B). The samples in lanes 1 and 21 were prepared from the colonic cell

25   line T84; lane 2 from fresh mouse salivary tissue; and lane 3 from untransfected LTK-cells. Lanes 4 to 7 contain RNA prepared from independent clones generated from pSTK7 transfection; lanes 8 to 15 from cotransfection with pSTK7 and pCOF-I; and lanes 16 to 20 from co-transfection with pSTK7 and pCONZ. The 28S and

9

18S rRNA bands are indicated. The position anticipated for the full-length transcript (6.2 kb) is also marked. The expected transcript for the HSV YK gene is 2.2 kb.

Figure 13 is Western Blot analysis. Total protein extracts were
5   prepared from the indicated cell lines, separated on a 7.5% polyacrylamide-SDS gel and transferred to nitrocellulose as described herein. Lane 1: extract from line 2a-4A; lane 2: 4a-2D; lane 3: 4a-3I; lane 4:6b-J. The blot was incubated with a mouse monoclonal antibody and followed by anti-mouse antibody. The immune complexes were visualized with alkaline phosphatase
10   as described herein. Protein molecular weight standards are shown on the left. The arrow on the right marks the 170 kilodalton band.

Figure 14 is expression vector pCOF-1. The complete CFTR coding region (open boxes) is positioned downstream from the human metallothionein IIa (hMTIIa) promoter (hatched box). The human
15   metallothionein IIa initiation codon is joined with that of CFTR at an Nco I site introduced by the synthetic oligonucleotides.
5'-CACTGCAGACCATGGAGAGGTCGCCTCTGGAAAAGGCCAGCGTT- 3' and 5'-GACTGCAGCTGAAAAAAAGTTTGGAGACAACGCTGGCCTT-3'. The DNA sequence from exons 2 to 24 and its 3' flanking region was derived from
20   clones 10-1 (Pvu II-Xba I)), TI6-1 (Xba I—Sph I), C1-1/5 (Sph I—BstXI), T16E4.5 (BstXI-Nco I), and TE27E2.3 (Nco I-Xho I), as indicated. In pCOFΔF508, the 1-kilobase (kb) BamIII-Hpa a fragment was replaced with the corresponding fragment from clone C1-1/5.

Figure 15 is RNA and protein analysis of mouse L cell lines expressing
25   human CFTR. (A) Total RNA from Caco2 (5 μg, lane 1), LTK (10 μg, lane 2), 6B-I (10 μg, lane 3), 4a-3I (10 μg, lane 4), and 5-2D (10 μg, lane 5) cell lines were electrophoresed on a 1% formaldehyde/agarose gel, transferred to Hybond™-N (Amersham), an hybridized with ^{32}P-labelled cDNA probe. The 6.2-kb CFTR mRNA of the Caco2 cell line is indicated with the arrow. The
30   relative positions of the 28S and

18S rRNAs as indicated (B) Protein fractions from nuclei and mitochondria (lanes A) and curde light membranes (lanes B) from the cell lines T84, LTK⁻, 6B-I, 4a-3I and 5-2D. Bands were visualized by ^{125}I-labelled rabbit anti-
5 mouse antibody. Molecular mass in kDa are indicated.

Figure 16 is video microscopic detection of Cl⁻ permeability in single L cells, SPQ fluorescence intensities (F) are expressed reltive to SPQ fluorescence intensity in the absence of Cl⁻ quenching (F_o; the
10 direction of changes in F_o/F reflect parallel changes in intracellular Cl⁻ concentration. (A-C) Single cells, each indicated by a different symbol, are shown containing a frameshift construct leading to a predicted truncated CFTR (A); a mutant construct CFTRΔF508 (B); and an intact
15 construct, CFTR (C). Forskolin (10 μM) and NO$_3^-$ medium were perfushed over the cells for the periods indicated by solid bars. Time scale in C applies to A-C. (D) A single CFTR cell was repeatedly pulsed with NO$_3^-$ medium. Rate of fluorescence chage as well as peak response
20 during constant pulses (1 min) was unchanged over the time course of exposure to forskolin, indicating that the cAMP-induced Cl⁻ permeability is sustained. (E) Simultaneous determinations of Cl⁻ concentrations (F_o/F) (open and solid squares) and cell volume, expressed as
25 relative area of constant optical section (open and solid circles), in a single cell expressing CFTR (solid squares and circles)) and a single cell expressing truncated CFTR (open squares and circles). As indicated, the cells were exposed to 5 μM gramicidin and to 10 μM forskolin.
30 Figure 17 is whole-cell Cl⁻ currents in transfected L cell. (A) Time course of whole-cell currents measured from cells transfected with the frame-shift CFTR control construct (solid diamonds) and the intact CFTR construct (open squares) after the addition of a solution
35 containing 10 μM forskolin, 1 mM isobutylmethylxanthine, and 100 μM N^6, O$^{2'}$-dibutylyladenosine 3',5'-cyclic monophosphate. The arrow indicates the time of solution

addition. Outward currents were measured at E_m = +20 mV. (B) Whole cell
current-voltage relationships for a CFTR-expressing cell before (Left) and 3

5   min after (Right) induced activation. The current scale for the nonstimulated
cells is shown enlarged as indicated. The dashed line indicates the zero-
current level. (C) Mean current-voltage relationships for CFTR cells (n = 10)
before and after activation (solid circles and open squares, respectively).
Replacement of bath NaC1 with sodium gluconate resulted in a shift in

10   reversal protention of cAMP-activated currents (open triangle) (n = 4).

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The normal CFTR gene sequence and protein sequence is shown in
Figure 1. The sequence consists of cDNA identified and isolated from

15   chromosome of mammalian cells in accordance with the procedure identified
in United States Patent No. 5,981,178. It is believed that the cDNA of CFTR
is made up of at least 27 exons as identified in applicant's pending Canadian
patent application serial number 2,066,204 published March 7, 1991.

The exact cause of the sequence instability is unknown, but the region

20   frequently involved in this rearrangement is a 13 bp direct repeat in exon 6a of
the CFTR gene as shown in Figure 3. This region also shows sequence
identity with consensus regulatory sequences in E. coli as noted in Figure 3. It
is possible that transcription initiates from this sequence, resulting in a product
which is toxic to the bacterial host. It is also possible that the DNA sequence

25   itself elicits an inhibitory response. In consequence, the transformation
frequency and growth of the desired plasmid construct become extremely
poor. The rearrangement of this sequence detected in the recovered plasmids
may have alleviated its toxicity in bacteria. Thus, the inability to obtain a full-
length CFTR cDNA in the original cloning experiment (Rommens et al, 1989)

30   may also be explained.

Since most of the spontaneous sequence
rearrangements detected disrupts the open reading frame
of the full-length cDNA, none of the resulting clones are
expected to be useful in functional complementation
5   studies. Although it is possible to use alternative
host/vector systems to overcome this problem, the
convenience of the E. coli/pBR322-based cloning system
required a better system.

The region of interest with respect to enhancing
10   propagation of the cDNA in microorganisms and in
particular in E. coli is in exon 6. A 13 base pair repeat
has been found in exon 6. With reference to Figure 1, the
region of exon 6 is identified. More specifically,
within exon 6 the repeat sequences occur at positions 923
15   through 935 and at 981 through 993. These repeat
sequences are separated by approximately 44 base pairs.

It has been discovered that by modifying one or more
of the base pairs in region 924 through 936 or 982
through 994 without changing the amino acid encoded by
20   the respective codon results quite surprisingly in the
propagation of the cDNA in a suitable microorganism, such
as E. coli. The modified cDNA, in one of the repeat base
pairs sections, does not produce a toxic effect in the
microorganism so that a stable propagation of the cDNA
25   can proceed. The benefit in the stable propagation of
cDNA is not only from the standpoint of expression of the
gene to produce the protein, but also for producing
multiple copies of the cDNA for the CFTR gene to enable
other uses of the gene, such as in drug and gene therapy
30   for correcting the effects of CF.

Construction of Full Length cDNA

Two types of full length cDNA clones have been
constructed to evaluate the modified cDNA. One of them
is for efficient propagation in E. coli and the other one
35   for expression in mammalian cells.

Since the entire coding region of CFTR is present in
multiple overlapping clones, it is necessary to construct

13

the full-length cDNA in three separate steps (Figure 4):
construction of the 5'end of the cDNA (from the beginning
of the coding region to the XbaI site in exon 5),
(Rommens et al, 1989) generation of the middle segment
5   (from the XBaI site to the SphI site in exon 10)
including the critical region by the polymerase chain
reaction (PCR: Saiki et al, 1985), and (Riordan et al,
1989) construction of the 3'end of the cDNA from
different existing clones (from SphI site to exon 24).
10 The starting clones of this work have been described
previously (Riordan et al, 1989).

        To construct the 5'end of the cDNA in the mammalian
expression vector, oligonucleotides were used to
synthesize the open reading frame between the initiation
15 codon and the end of exon 1 (the PvuII site). This
fragment was inserted between the promotor of vector
pSG3X and exon 2 of cDNA clone 10-1 (see Figure 5). A
single nucleotide modification (C to G) was introduced at
the position immediately after the initiation codon to
20 facilitate subsequent manipulation. The alteration also
introduced an amino acid change from Glutamine to
Glutamate.

        To construct the 3' end of the cDNA, DNA fragments
were obtained from three existing clones (Figure 6): SphI
25 to BstXI (exons 10-20) from C1-1/5 (Riordan et al, 1989),
BstXI to NcoI (exons 20-24) from T16-4.5 (Riordan et al,
1989), and NcoI to EcoRI (exon 24 to about 100 bp
downstream from the polyadenylation signal) from TE27
(Rommens et al, 1989). The genomic DNA fragment included
30 at the end of this construct to ensure proper
transcription termination and mRNA processing. This
3'end construct was used for both the expression and
propagation vectors.

        The cDNA clone T16-1 (Riordan et al, 1989) was used
35 as the template for generation of the middle segment,
which was used for both vector systems. To introduce a
sequence modification in one of the 13 bp repeat region

14

of exon 6, the _in vitro_ mutagenesis procedure based
oligonucleotide-directed PCR (Higuchi et al, 1988; Ho et
al, 1989) was used. Briefly, two overlapping segments,
spanning exons 3 to 6 (536 bp) and exon 6 to 17 (867 bp),
5   were generated from T16-1 by PCR with two sets of
oligonucleosides (see Figure 7). The two overlapping
segments were joined by the PCR procedure with the
outermost, flanking oligonucleotide primers. The
product, 1.38 kb in size, was then digested with XbaI and
10 SphI to generate the modified middle segment.

The pBluescript® vector from Stratagene was used as
the basis propagation vector. The 5' end of the full-
length cDNA in this vector was derived from cDNA clone
10-1 (between the PstI and XbaI sites). The resulting
15 clones were named pBQ4.7 and pBQ6.2 (Figure 8).

The plasmid pSGM3X was used as the basic mammalian
expression vector. It is similar to p SGM1 previously
reported by Meakin et al (1987), except that the human
metallothionein promoter was inserted in an opposite
20 orientation and that a XhoI site was inserted in the KpnI
site within the Ecogpt gene. The resulting expression
vector was named pCOF-1 (Figure 5).

Sequence Verification

Since errors are often introduced into PCR products
25 due to infidelity of the Taq polymerase, the middle
segment of the cDNA clones were first verified by DNA
sequencing. The procedure of dideoxy termination was
used according to the US Biochemical Sequenase® kit.
Consequently, it was necessary to replace a small segment
30 (the XbaI site in exon 6 to the FspI site in exon 7) of
the clone with the corresponding segment from T1601.
Overlapping DNA sequencing was then performed for the
entire open reading frame of each of the completed full-
length cDNA constructs.

35 **In vitro Translation**

To produce large quantities of RNA from the
propagation vector (pBQ4.7 and pBQ6.2) for the purpose of

in vitro translation, the mCAP® mRNA capping kit from Stratagene was used.

Since the open reading frame predicts a protein of 170,999 kilodaltons (kd), it would be difficult to
5 produce a product of this size _in vitro_. It was therefore necessary to use differently prepared templates to examine various segments of the open reading frame. To prepare these templates, the propagation vector pBQ6.2) was digested with FspI, EcoRI, HpaI and XhoI,
10 accordingly.

in vitro translation was performed with a rabbit reticulocyte lysate kit from Promega Corporation and ^{14}C-labelled leucine from Amersham.

Transfection Studies in Mouse LTK Cells
15 The calcium phosphate co-precipitation procedure for introduction of plasmid DNA into mouse LTK-cells has been described previously (Meakin et al, 1987). The plasmid pSTK7 was used for cotransfection with the expression vector pCOF-1. As a control vector, a plasmid construct
20 similar to pCOF-1, except for the deletion of a single base pair in exon 1 was used; this deletion was expected to result in premature termination of translation.

The mouse LTK-cells were passaged in α-MEM medium supplemented with glutamine, 10% fetal bovine serum and
25 antibiotics. Biochemical selection for TK positive cells was achieved in medium containing hypoxanthine, aminopterin and thymidine (HAT medium).

DNA and RNA Analyses
Plasmid and DNA samples were prepared from the
30 bacterial cells and genomic DNA from transfected mouse L cells. Total RNA was extracted from mammalian cells for examination of gene expression. Standard procedures, essentially as described by Sambrook et al (1989), for DNA and RNA analyses were used.
35 **Protein Analysis**
Animal cells were harvested in TEN buffer (40 mM Tris-HCl, pH 7.5, 1 mM EDTA, 150 mM), after the cells

16

were washed in phosphate-buffered saline. Cells were collected by scraping, resuspended in 2560 mM Tris-HCl (pH 8), and stored in -80°C ready for total protein extraction. For SDS-polyacrylamide gel electrophoresis,
5 the frozen cells were resuspended in the loading buffer and boiled for 5 min prior to electrophoresis according to the procedure of Laemmli et al (1970).

Transfer of protein from polyacrylamide gel to nitrocellulose was performed overnight at 0.4 mA
10 according to the procedure of Towbin et al (1979). The membrane was baked at 80°C for 90 min. Detection of CFTR was accomplished with the use of a monoclonal antibody (gift of N. Kartner and J. Riordan) prepared against a portion of the predicted protein (amino acid residues
15 347-698) and the ProtoBlot® AP System according to procedures recommended by the supplier (Promega Corporation) with 1% BAS as a blocking agent.

Vector Construction

Three plasmid constructs were made in this study -
20 pBQ4.7; pBQ6.2 and pCOF-1 as described in the previous section. pBQ4.7 is a full-length cDNA clone contained in the Bluescript® vector; it contains the entire coding region except the 3' untranslated region. pBQ6.2 contains the same sequence as in pBW4.7 plus the 3' untranslated
25 region. Both of these plasmids contain modification in the first 13 bp direct repeat in exon 6b (Figure 3). The full-length cDNA construct in pCOF-1 is inserted downstream from the human metallothionein IIA promoter in the vector pSGM3X and the coding region is flanked by
30 CFTR genomic DNA sequence at the 3' end. In addition to the modification introduced into the exon 6b region of the cDNA, as in the two previous constructs, pCOF-1 has another modification (C to G) in the first nucleotide after the presumptive initiation (ATG) codon in exon 1.
35 As the first step in examining their integrity, restriction enzyme digestions were used to derive a map for each of these full-length cDNA constructs. The

17

results showed that all of them produced a set of DNA fragments as expected. It is, however, of interest to note that there was one DNA fragment in each of the digestions which appeared to have a different mobility

5   when compared to the pattern of the corresponding parental cDNA clones displayed on polyacrylamide gel. The mapping data located the fragments with the altered mobility to those containing the 13 bp direct repeats.

To ensure the DNA sequence was intact as designed,

10 the middle segment containing the modified segment was examined by DNA sequencing. The result showed that 3 bp were altered in the region. While two of the modifications (T to C at position 930 and A to G at 933) were inserted as expected, an additional change (T to C

15 at position 936) was found (Figure 9). This unanticipated alteration was found in all five clones examined, suggesting that it was introduced by an error in the synthesis of the oligonucleotide primer. Nevertheless, this error was located outside of the "-35"

20 prokaryotic consensus sequence (Hawley and McClure 1983) and did not alter the encoded amino acid (Figure 3).

Two other undesired nucleotide changes were detected in the middle cDNA segment. Since these latter substitutions, probably introduced during PCR, would

25 change the encoded amino acids, it was necessary to replace the region containing these changes with a segment from the original plasmid.

The DNA sequence in all three resulting clones appeared to be stable after long term propagation in E.

30 coli. Their entire coding region was sequenced and no additional changes were detected. It was noted that the transformation efficiency of these constructs in bacteria increased substantially (100 times higher than the unmodified cDNA) and that the copy number of plasmids (as

35 reflected by the yield in DNA preparation) also increased.

2091907

18

Protein Synthesis by in vitro Translation

To ensure that the open reading in the full-length cDNA was uninterrupted, the _in vitro_ translation method was performed. Accordingly, the propagation vector was
5 linearized at an appropriate restriction site in the circular plasmid and used as the template for production of RNA suitable for translation _in vitro_. Since the entire CFTR protein would be probably too large to be translated _in vitro_, the experiment also included shorter
10 transcripts produced from templates interrupted within the coding region.

An example of the result is shown in Figure 10. The template used for this experiment was pBQ6.2 linearized at the HpaI site; the _in vitro_ translation product was
15 expected to be 86.2 kd. As shown in Figure 10, a band migrating at the position of the expected molecular weight is clearly visible, indicating that translation initiates at the ATG codon as predicted and continues through the HpaI site in exon 13. The identity of this
20 protein is also confirmed by its ability to react with a monoclonal antibody against CFTR. The other, smaller proteins observed in the products are probably results of premature termination or intenal initiation of translation.

25 Similar experiments were performed with the plasmid treated with XbaI and EcoRI. Protein bands corresponding to the expected translation products 19 and 77.5 kd, respectively, were readily detectable on the polyacrylamide gel. These results, therefore, provided
30 further confirmation that the predicted open reading frame was correct and intact in the full-length cDNA constructs.

Expression in Mouse L Cells

To examine if CFTR could be produced in heterologous
35 mammalian cells, the pCOF-1 plasmid was used to transfect mouse LTK-cells. Another plasmid containing the herpes simplex virus TK gene was included in the transfection to

allow biochemical selection of cells that were
successfully transfected. A plasmid (pCONZ) which
suffered a single base pair deletion immediately adjacent
to the 3' initiation codon was used in a parallel
5   transfection experiment as a negative control (as the
frameshift would result in no CFTR products).
DNA Analysis

HAT-resistant L cell colonies were isolated and
expanded into individual mass cultures and genomic DNA
10 were isolated from these cultures for characterization of
integrated plasmid DNA. As expected, all of the HAT-
resistant L cell lines was found to contain an intact HSV
TK gene (as demonstrated by gel-blot hybridization
analysis). In addition, the copy number for the
15 integrated plasmid DNA was found to vary among different
lines, an anticipated from the calcium phosphate co-
precipitation protocol. Gel-blot hybridization analysis
was then performed for the cell cultures transfected with
pCOF-1 and pCONZ. The full-length cDNA clone (a 6.2 kb
20 PstI fragment from pBQ6.2) was used as proof to examine
the CFTR sequence.

Among the 15 pCOF-1/pSTK7 co-transfected cell lines,
14 were found to contain at least a portion of the CFTR
cDNA, as judged by the hybridization pattern of EcoRI-
25 digested DNA from these lines. Five of these lines shows
a 1.5 kb and 2.5 kb band, as predicated for an intact
cDNA. For the remaining lines, the size of the EcoRI
fragment was different from those predicted, indicating
rearrangement or integration occurred through this region
30 of the cDNA. Upon further analysis with BamHI and NcoI
digestion, however, only four of the five lines showed
intact 5' and 3' ends of the cDNA (see Figure 11). The
apparently intact clones were 4a-2C, 4a-3I, 4a-3K, and
4a-4S.

35   Similar DNA analysis was performed for cell lines
co-transfected with pCONZ and pSTK7. Among the 10 clones
examined, nine appeared to contain CFTR sequence and

eight of them showed the predicted 1.5 kb and 2.5 kb EcorI fragments (see Figure 12). Further analysis with BamHI and NcoI on five of them revealed that four (6a-1D; 6a-2F; 6b-J and 6b-K) contained intact 5' and 3' ends.

5   **RNA Analysis**

Expression of CFTR in the transformed mouse cell lines was next examined by RNA blot hybridization analysis. Two of the pCOF-1 transfected lines (4a-3I and 4a-3K) were found to express high levels of RNA (see

10   Figure 12a) but the remaining two lines 4a-1C and 4a-4Q) were low. In addition to the 6.2 kb band expected for correctly initiated and terminated CFTR RNA in these lines; however, hybridizing sequences were also detected at the 9-10 kb range, suggesting improper initiation or

15   termination. A similar, high expression pattern was observed for three of the lines transfected with pCONZ (6a-1A; 6a-1D and 6b-J) (Figure 12a). For clone 6a-2F, only a low level of the 6.2 kb species was detected. No RNA could be found for clone 6b-K, despite the presence

20   of an apparent intact CFTR sequence.

The level of CFTR-derived RNA present varied among the cell lines mentioned above, with the strongest signals seen in clones 4a-3I and 6b-J. The variation was not due to poor sample preparation, as confirmed by

25   hybridization analysis of the same blot with the TK probe (Figure 12B).

Protein Analysis

Total protein was then extracted from several representative cell lines and examined for the presence

30   of CFTR-related products. The interpretation of the results of this experiment was complicated by the presence of cross-reactive materials present in the mouse L cell, although no CFTR-related RNA was detectable. Figure 13 is an illustration in drawing format of the

35   Western blot results. These results are based on a monoclonal antibody (L12B4) prepared against CFTR which is able to detect faint bands (at regions 170, 90, and 74

21

kd) in all cell extracts. Nevertheless, the intensity of
the 170 kd band, the one expected for the CFTR protein,
was clearly much stronger in 4a-3I, the cell line
containing an intact full-length CFTR cDNA. Only
5 background bands were detected for 2a-4A (containing
pSTK7 only), 4a-2D (incomplete pCOF-1 sequence), and 6b-J
(pCONZ). The presence of the CFTR protein in 4a-3I was
also confirmed in partially purified membrane fractions.

1) A single 170kdalton band was observed in the
10 extract from the cell line 4a-3I. This cell had been
demonstrated by DNA analysis to contain multiple copies
of the intact promoter of the expression vector and of
the CFTR gene. RNA analysis also indicated expression of
high levels of mRNA for this clone.

15   2) In contrast, the cell line 6b-J, generated with
the mutant gene does not express a reacting 170kdalton
band. This line generated by co-transfection with the
pCONZ plasmid was shown to have an intact promoter and
gene portion. In addition, high levels of RNA were
20 detected. Identical results for these samples were
obtained even when the extracts were not subjected to
heating prior to gel loading.

3) 170kdalton bands were not observed in the cell
line 4a-2D generated by the transfection experiment with
25 the pCOF-1 plasmid and the cell line 2a-4A generated by
the transfection with the pTK7 plasmid only. Genomic DNA
analysis of the cell line 4a-2D indicated that only
portions of the CFTR gene had been incorporated. As
expected, no CFTR RNA could be detected.

30   In conclusion, the cell line 4a-3I appears to be
producing a 170kdalton CFTR protein as expected from the
open reading frame of the pCOF-1 plasmid. For
comparison, also shown on the protein analysis figure are
crude membrane preparations of the cell lines Panc-1 (a
35 negative control) and T84 (provided by N. Kartner and J.
Riordan). Multiple bands are observed in the T84
preparation. These differences suggest that post-

translational modification including glycosylation and/or phosphorylations have occurred to the CFTR protein in T84 cells.

5   A CAMP-inducible chloride permeability has been detected in mouse fibroblast (L cell) lines upon stable integration of a full-length cDNA encoding the human cystic fibrosis transmembrane conductance regulator (CFTR). As indicated by a Cl^--indicator dye, the Cl^- permeability of the plasma membrane increases by 10-to

10   30-fold within 2 min after treatment of the cells with forskolin, an activator of adenylyl cyclase. The properties of the conductance are similar to those described in secretory epithelial cells; the whole-cell current-voltage relationship is linear and there is no

15   evidence of voltage-dependent inactivation or activation. In contrast, this cAMP-dependent Cl^- flux is undetectable in the untransected cells or cells harboring defective cDNA constructs, including one with a phenylalanine deletion at amino acid position 508 (ΔF508), the most

20   common mutation causing cystic fibrosis. Those observations are consistent with the hypothesis that the CFTR is a cAMP dependent Cl^- channel. The availability of a heterologous (nonepithelial) cell type expressing the CFTR offers an excellent system to understand the basic

25   mechanisms underlying this CFTR-associated ion permeability and to study the structure and function of the CFTR.

We describe the construction and use of a mammalian expression vector to produce human CFTR in a long-term

30   mouse fibroblast culture. We show that expression of CFTR induces a cAMP-dependent Cl^- conductance, which is normally not observed in these cells. This expression system is suitable for study of the Cl^- conductance pathway and its regulation and to provide medical

35   treatment for CF.

Plasmid Vectors

The mammalian expression vector pCOF-1 is a derivative of pSGM3X, which is similar to pSMG1 (Meakin et al), except that the human metallothionein IIa

5 promoter (Karin et al, 1982) was inserted in the opposite orientation and a Xho I site was inserted in the Kpn I site within the Ecogpt gene.   To reconstruct the full-length CFTR cDNA in pCOF-1 (Fig. 14), the bulk of the coding region (exons 2-24) was obtained from partial

10 cDNA clones (Riordan et al, 1989), except that the three silent nucleotide substitutions (T → C at position 930, A → G at position 933, and T → C at position 936) were introduced into the exon 6b region with oligonucleotide-mediated mutagenesis by the polymerase chain reaction

15 (19, 20). The 3' untranslated region of CFTR in pCOF-1 was derived from the genomic DNA clone TE27E2.3 (Rommens et al, 1989). The entire coding region of exon 1 (from the initiation codon to the Pvu II site) was generated by two complementary synthetic oligonucleotides and the

20 Klenow fragment of DNA polymerase I, where a single nucleotide substitution (C → G) was introduced immediately after the initiation codon (underlined in the legend to Fig. 14) to create a Nco I site for ligation to the human metallothionein IIa promoter. The latter

25 substitution changed the encoded amino acid glutamine to glutamic acid. The construction of control plasmid pCONZ was similar to that of pCOF-1, except that a single nucleotide was deleted 35 base pairs downstream from the initiation codon. A truncated protein would be predicted

30 from this frameshift construct. The plasmid pCOFΔF508 was generated by replacing sequences of exons 9-13 in pCOF-1 with the corresponding fragment from C1-1/5, a cDNA containing the ΔF508 mutation. The full-length cDNA clone pBQ6.2 contained a 6.2 kb Pst I fragment in

35 pBluescript (Stratagene) and was constructed similarly to pCOF-1 except that the exon 1 region was derived from clone 10-1. The integrity of the CFTR cDNA inserts in

pBQ6.2, pCOF-1, and the critical regions in the other plasmid constructs were verified by DNA sequencing. The plasmid vector (pSTK7) containing the herpes simplex virus thymidine kinase gene, used in the cotransfection experiments, has been described (Meakin et al). Bacterial cell cultures and plasmid DNA samples were prepared according to standard procedures (Sambrook et al, 1989).

Cell Culture and DNA Transfection

Each of the three test plasmids (20 μg), pCOF-1, pCONZ, and pCOFΔF508, was cotransfected with pSTK7 (1 μg) into mouse LTK$^-$ cells by calcium phosphate coprecipitation (Meakin et al). Biochemical selection for thymidine kinase-positive cell was achieved in minimal essential medium supplemented with hypoxanthine/aminopterin/thymidine (HAT;gibco/brl). In some experiments, the test plasmids were linearized at the unique _Sfi_ I site (Fig. 14). High molecular weight DNA was isolated from each clonal cell line (Miller et al, 1988), digested with restriction enzymes EcoRI, BamHI, and Nco I, and analyzed by agarose-gel-blot hybridization (Sambrook et al, 1989) with the full-length cDNA (insert from pBQ6.2) as probe. Total RNA was extracted (MacDonald et al, 1987) and analyzed by agarose-gel-blot hybridization (Sambrook et al, 1989).

Protein Analysis

Cells were homogenized in a hypotonic buffer containing 10 mM KCl, 1.5 mM MgCl$_2$, and 10 mM Tris·HCl (pH 7.4). Nucleic and mitochodria were collected by centrifugation at 4000 x g for 5 min (fraction A). A crude membrane fraction was then collected by centrifugation at 9000 x g for 15 min (fraction B). Membrane pellets were dissolved in loading buffer and separated on a NaDodSO$_4$/polyacrylamide (6%) gel (Laemmli, 1970). Proteins were transferred to nitrocellulose as described (Towbin et al, 1979) and immunodetected with monoclonal antibody M3A7 (Kartner et al, 1991).

6-Methoxy-1-(3-sulfonatopropyl)quinolinium (SPGQ) Fluorescence Assay

L cells grown on glass coverslips for 1-2 days were uniformly loaded with the Cl⁻-indicator dye SPQ by
5   incubation in hypotonic (1:2 dilution) medium containing 20 mM SPQ at room temperature for 4 min. The mounted coverslip was perfused continuously at room temperature with medium containing 138 mM NaCl, 2.4 mM K₂HPO₄, 0.8 mM KH₂PO₄, 10 mM Hepes, 1 mM CaCl₂, 10 mM glucose, and 10 μM
10   bumetanide (pH 7.4) on the stage of an inverted microscope. NO₃⁻ medium was identical except that NO₃⁻ replaces all but 10 mM Cl⁻. To minimize Cl⁻ fluxes through nonconductive pathways, we performed the experiments in the absence of HCO₃⁻ and in the presence of
15   bumetanine, inhibiting the anion exchanger and Cl⁻/cation cotransporters, respectively. Fluorescence and the differential interference contrast imagine were performed simultaneously (25-27). SPQ fluorescence intensities (F) were normalized to total SPQ fluorescence F_o, determined
20   as F measured in the absence of intracellular Cl⁻, since autofluorescence was negligible, Calibration (n = 7 cells) were performed essentially as described (Foskett, 1990). The effective quenching constant K_{cl} was 15 M⁻¹; the testing intracellular Cl⁻ concentration was ≈ 70 mM.
25   Cell volume changes were obtained by planimetry of differential interference contrast images (26, 27) and are presented as relative changes in the areas of the measured optical sections. By exposure of the cells to media of various osmolarities, we observed that
30   differential interference contrast imaging of a single optical section can detect volume changes.

Whole-Cell Current Recordings

Membrane currents were measured at room temperature 12-24 hr after plating cells with whole-cell patch-clamp
35   techniques (Hamill et al, 1981). The patch pipet contained 110 mM sodium gluconate, 25 mM NaCl, 8 mM MgCl₂, 10 mM hepes, 4 mM Na₂ATP, and 5 mM Na₂EGTA (pH 7.2). The

bath contained 135 mM NaCl, 2.4 mM K_2HPO_4, 0.8 mM DH_2PO_4, 3 mM $MgCl_2$, 1 mM $CaCl_2$, 10 mM Hepes, and 10 mM glucose (pH 7.2). To examine the time course of a cAMP-evoked conductance changes, membrane potentials were alternately
5   clamped at -30 and +20 mV for 600 and 400 ms, respectively. Current-voltage relationships were determined by measuring the currents at the end of 400 - ms voltage steps from 0 mV to ± 70 mV (10 mV increments). Cell capacitance was compensated using cancellation
10   circuitry of the EPC7 patch-clamp amplifier.

DNA Transfection

Stable mouse fibroblast cell lines containing full-length CFTR or mutant cDNA, as well as cell lines with the pSTK7 plasmid alone, were established. Four cell
15   lines with pCOF-1 (4a-2C, 4a-3I, 4a-3K, and 4a-S) contained an intact human metallothionein IIa promoter and the CFTR coding region, as judged by DNA analysis (data not shown). Similarly, four cell lines (6a-1D, 6a-2F, 6b-I, and 6b-K) for pCONZ control cDNA with the
20   frame-shift mutation three for pCOFΔF508 (5-2C, 5-1A, and 5-2D) cDNA with the major CF mutation and two for pSTK7 (2a-4A and 2a-3C) were identified).

RNA Analysis

Abundant levels of RNA transcripts of ≈ 10 kb in
25   size were detected in cDNA transected cells Fig. 15A. Although their size was larger than the anticipated 6.2 kb, the result appeared to be consistent among all cell lines. It seemed probable that an alternative polyadenylylation site(s) instead of those contained in
30   the cDNA constructs were utilized. Thymidine kinase-specific transcripts were detected in all HAT -resistant cell lines tested.

Protein Analysis

Cells expressing CFTR mRNA (as represented by the
35   cell line 4a-3I) contained an antibody-reacting protein band that was indistinguishable from mature CFTR expressed endogenously in membranes of the colonic

carcinoma cell line T-84 (Fig. 15B). The amount of
protein was within the range of that observed for T-84
cells, with a significant portion in the light membrane
fraction. Reacting bands were not detected in

5 untransfected LTK⁻ cells or in cell transfected with CFTR
gene predicted to produce a truncated product (line
6B-I). The latter result was expected as the antibody
was directed against the C terminus of the protein. No
immunoreactive material was observed for CFTRΔF508 in the

10 examined fractions (line 5-2D) (Cheng et al, 1990). It
was therefore uncertain if the mutated protein was
produced in these transfected cells.

SPQ Fluorescence Assay

To investigate if expression of CFTR affected Cl⁻

15 conductance, a single-cell assay based on quantitative
fluorescence intensity measurements of the Cl⁻ indicator
dye SPQ was performed. The basic protocol involved
exposing cells to NO₃⁻ medium followed by return to normal
Cl⁻ medium. Since NO₃⁻ is generally permeable through Cl⁻

20 channels but, unlike Cl⁻, does not quench spq, CHANGES IN
spq fluorescence intensities due to these anion
substitutions measure unidirectional Cl⁻ fluxes; the rate
of change measures cell Cl⁻ permeability. 10 μM forskolin
was added subsequently to increase intracellular levels

25 of cAMP. After a 2-min exposure to forskolin, the medium
was again switched to the NO₃⁻ medium, in the continued
presence of forskolin. Thus each cell was used as its
own control to evaluate the Cl⁻ permeability induced by
cAMP.

30 Only slow changes in SPQ fluorescence intensity
were observed in cells with the frame-shift DFTR
construct (as represented) by line 6B-I) upon exposure to
NO₃⁻ medium (Fig. 16A), indicating that these cells
maintained a low resting Cl⁻ permeability. Exposure of

35 these transfected control cells to forskolin did not have
any effect (n = 41 cells from four passages). Similar
results were obtained from cells transfected with the

thymidine kinase gene only (line 2a-4A, n = 22 cells) and from untransfected cells (n = 18 cells).

Exposure of cells expressing CFTR to NO_3^- medium similarly elicited little or no change in SPQ fluorescence intensity (as represented by line 4a-3I in Fig. 16C), indicating that CFTR expression per se did not enhance Cl⁻ permeability. In contrast, a second exposure to NO_3^-, during forskolin stimulation, caused a rapid loss of intracellular Cl⁻ (Fig. 16C). This response was highly reproducible; there was a 20- to 30-fold increase of unidirectional Cl⁻ flux (Illsley et al, 1987) from ≈ 0.03 mM/s to 0.9 mM/s for each of the 10 cells in the microscopic field (Fig. 16C). Of 106 cells examined from eight passages, all responded similarly. Similar results were obtained from another cell line expressing CFTR (clone 4a-3K, n = 23 cells). In the continued presence of forskolin (Fig. 16D), Cl⁻ permeability remained enhances at near maximal levels for as long as 30 min (n = 12 cells).

There were no volume changes during exposure of forskolin-stimulated CFTR cells to NO_3^-, indicating that the substantial changes in the intracellular concentration of Cl⁻ were not associated with changes in cell salt content. Therefore, the Cl⁻ fluxes were likely to be associated with equal fluxes of NO_3^- in the opposite direction. The lack of changes in intracellular Cl⁻ concentration or cell volume during forskolin stimulation in Cl⁻ medium demonstrates that neither CFTR expression nor cAMP conferred enhanced cation conductance. Together with the enhanced Cl⁻ permeability, cell shrinkage or swelling would have been observed if there were a high K^+ conductance or high Na^+ conductance, respectively.

Cl⁻ permeability was also examined in cells containing a CFTRΔF508 construct. These cells (as represented by clone 5-2D, n = 15 cells) exhibited low Cl⁻ permeability under resting conditions and the

permeability could not be enhanced by forskolin (Fig. 16B).

To establish that the cAMP-induced Cl⁻ permeability in the CFTR-expressing cells was due to activation of a
5   Cl⁻ conductance, gramicidin was included in the normal Cl⁻ medium to increase cation conductance of the plasma membrane. Under these conditions, the presence of Cl⁻ channel would result in a substantial influx of both Na⁺ and Cl⁻, causing extensive cell swelling. Exposure to
10   gramicidin had no effect on SPQ fluorescence or cell volume in cells with the frame-shift CFTR construct (n = 13 cells) or an intact CFTR construct (n = 33 cells) (Fig. 16E), supporting our conclusion (above) that resting Cl⁻ conductance was negligible in both the control
15   and CFTR cells. After the addition of forskolin, however, there was a marked rapid cell swelling, after a lag period of from 30 to 90 s, accompanied by elevated intracellular Cl⁻ concentration in the cells expressing CFTR but not in control cells (Fig. 16E). These results
20   demonstrated that the basis of cAMP-induced Cl⁻ permeability observed in CFTR-expressing cells was a Cl⁻ channel.

cAMP-Stimulated Cl⁻ Currents in the CFTR-Expressing Cells

Corroborating the results of the fluorescence assay,
25   whole-cell current was stimulated in cells expressing CFTR (Fig. 17A). After a lag period of ≈ 30 s, outward current in 11 of 11 cells increased dramatically to a plateau, which was sustained for 2-5 min before decreasing toward control values. This "run down"
30   contrasts with findings using the SPQ assay, possibly reflecting a depletion of cytosolic factors necessary for sustained activation in the whole-cell patch-clamp configuration. currents evoked by the activation cocktail did not display any time-dependent voltage
35   effects (Fig. 17B) contrast to the expressing cells, none of the 9 cells containing the frame-shift CFTR construct

exhibited a response to the activation cocktail (Fig. 17A).

Current-voltage relations were essentially linear in both cell transfected with the intact CFTR and the frame-
5 shift CFTR constructs (now shown). Slope conductances of 1.8 ± 0.4 nS and 1.9 ± 0.5 nS (n = 9) were calculated in control cells before and after treatment with the activation mixture, respectively. The slope conductance in CFTR-expressing cells was similar (1.6 ± 0.1 nS) but
10 treatment with the activation mixture induced an $\approx$ 13-fold increase to 20.1 ± 1.7 nS (n = 10) (Fig. 17C).

The reversal potential of the cAMP-activated current in cells expressing CFTR was -17 ± 5 mV, approaching the equilibrium potential for Cl⁻ under standard conditions
15 ($E_{Cl} = -32$ mV). From this measurement, the calculated anion versus cation permeability was $\approx$ 5:1. Anion selectivity was assessed in cAMP-activated cells by replacing bath NaCl (135 mM) with sodium gluconate (135 mM). This manipulation resulted in a shift in reversal
20 potential to $+33 \pm 8$ mV (n = 4), toward the predicted E_{Cl} of $+41$ mV (Fig. 17C).

The inability to construct a full-length CFTR cDNA has hampered progress in understanding the structure and function of the protein. The difficulty is mainly due to
25 the instability of the full-length sequence in E. coli, where there was sequence rearrangement often associated with a short (13 bp) direct repeat in exon 6b. Through the modification of DNA sequence within the first copy of the repeat, however, we have succeeded in the
30 construction of three different plasmids each containing the entire coding sequence. These plasmids replicate efficiently in E. coli (DH5α), and are intact after prolonged propagation. Although this particular modification involves the first repeat, it is understood
35 that this second repeat can also be similarly modified instead of, or in addition to, the first repeat to

achieve similar success in the propagation of the CFTR cDNA.

In addition to verification by direct DNA sequencing, the plasmids have also been examined for their ability to product proteins of expected sized _in vitro_ and _in vivo_. Based on the longest open reading frame of the consensus cDNA sequence (Riordan et al 1989), a protein of 170 kd (1480 amino acids) has been predicted as the CFTR gene product. The fact that a 170 kd band was detectable in the product translated _in vitro_ and in cells transfected with the full-length CFTR cDNA confirms the original prediction.

The availability of a full-length CFTR cDNA that can be expressed in mammalian and non-mammalian cells offers the opportunity to perform a detailed structure and function analysis of CFTR. The vectors described here are excellent tools for this purpose. With appropriate regulatory sequences inserted upstream of the coding region, it is understood that large quantities of CFTR may be produced through different kinds of heterologous gene expression systems, whereby various biochemical and biophysical studies can be performed.

The ability to express the full-length cDNA also allows development of functional assays for CFTR. In this context, Drumm et al (1990) have demonstrated through a retrovirus-intermediate that the modified full-length cDNA described here was able to confer the function of CFTR in a pancreatic carcinoma cell line (CPFAC-1) derived from a CF patient. Upon proper expression of the cDNA, the cAMP-mediated chloride transport activity was restored in this cell line, providing the first example of functional complementation of CFTR activity. The ability to confer CFTR expression in heterologous cells is an important step towards the possibility of gene therapy in the lung and pancreas of CF patients.

In order to understand the role of individual amino acid residues as well as regions of the protein, site-directed mutagenesis may be used to introduce additional mutations into the coding region of CFTR. The functional
5   assay may also be used to confirm if the sequence alterations detected in CF patients are bona fide disease-causing mutations. The latter consideration is important if broad scale disease diagnosis and carrier screening based on DNA information are to be implemented.
10   Further, since there is a general lack of genotypes, permanent cell lines in which the CF phenotype resists, the ability to generate heterologous cell lines capable of expressing various defective CFTR offers an alternative approach in understanding the function of
15   CFTR and in development of rational therapy. In the latter regard, procedures can be devised for screening of compounds that would interact with the defective protein and restore its function.

   To introduce additional mutations into the coding
20   region of CFTR, it is possible to replace regions of the cDNA with altered sequences, as demonstrated by the examples described above. The procedure is difficult, however, because many of the restriction enzyme sites involved are present in more than one vector. For this
25   reason, it is desirable to include unique restriction sites in the coding region of the cDNA. For example, by introducing a silent change (T to C) at position 1944 in exon 13, a novel BstEII site is created at the end of the sequence corresponding to the first NBF1 is probably the
30   most interesting region in CFTR because about one-third of the disease-causing mutations reside in the region. In combination with another unique site, such as SphI, it becomes extremely easy to replace sequences for NBF1.

   Finally, the full-length cDNA construct contained in
35   the pBQ4.7 and pBW6.2 may be excised in its entirety by a single PstI digestion or a double digestion with a combination of SaII, XhoI, SmaI or EcoRI. This

versatility allows the cDNA to be transferred from the current vector to other host-vector systems.

References

Cutting, G.R., Kasch, L.M., Rosenstein, B.J., Zielenski, J., Tsui, L.-C., Antonarakis., S.E. and Kazazian, H.H. Jr. (1990) "A Cluster of Cystic Fibrosis Mutations in the First Nucleotide-Binding Fold of the Cystic Fibrosis Conductance Regulator Protein" Nature 346:366-369.

Drumm, M.L., Pope, H.A., Cliff, W.H., Rommens, J.M., Marvin, S.A., Tsui, L.-C., Collins, F.S., Frizzel, R.A. and Wilson, J.M. (1990) "Correction of the Cystic Fibrosis Defect in vitro by Retrovirus-Mediated Gene Transfer" Cell 62:1227-1233

Hawley, D.K. and McClure, W.R., (1983) Nucleic Acids Res. 11:2237-2255

Higuchi, R., Krummel, B., and Saiki, R.K., (1988) "A General Method of in vitro Preparation and Specific Mutagenesis of DNA Fragments: Study of Protein and DNA Interactions", Nucleic Acids Res. 16:7351-7363.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989) "Site-Directed Mutagenesis by Overlap Extension Using the Polymerase Chain Reaction", Gene 77:51-59.

Kerem, B., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M. and Tsui, L.-C., (1989) "Identification of the Cystic Fibrosis Gene: Genetic Analysis" Science 245:1073-1080.

Laemmli, U.K., (1970) Nature 227:680-685.

Meakin, S.O., Du, R.P., Tsui, L.-C. and Breitman, M.L. (1987) "γ-Crystallins of the Human Eye Lens: Expression Analysis of Five Members of the Gene Family", Molec. Cell, Biol. 7:2671-2679.

Riordan, J.R., Rommens, J.M., Kerem, B., Alon, N.,
Rozmahel, R., Grzelchak, K., Zielenski, J., Lok, S.,
Plavsic, N., Chou J.-L., Drumm, M.L., Iannuzzi, M.C.,
Collins, F.S., and Tsui, L.-C., (1989) "Identification of
5 the Cystic Fibrosis Gene: Cloning and Characterization
of Complementary DNA", Science 245:1055-1073.

Rommens, J.M., Iannuzzi, M.C., Kerem, B., Melmer, G.,
Drumm, M.L., Melmer, G., Dean, M., Rozmahel, R., Cole
10 J.L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M.,
Riordan, J.R., Tsui, L.-C., and Collins, F.S., (1989).
"Identification of the Cystic Fibrosis Gene: Chromosome
Walking and Jumping" Science 245:1059-1065

15 Saiki, R., Sambrook, J., Fritsch, E.F. and Maniatis, T.,
(1989) Molecular Cloning Second Edition, Cold Spring
Harbor Laboratory Press, NY.

Twobin, H., Staehelin, T. and Gordon J., (1979) Proc.
20 Natl. Acad. Sci USA 76:4350-4354.

Boat, T., Welsh, M.J. & Beaudet, A. (1989) in The
Metabolic Basis of Inherited Disease, eds. Scriver, C.R.,
Beaudet, A.L., Sly. W.S. & Valle, D. (McGraw, Hill, New
25 York), 6th Ed., pp. 2649-2680.

Rommens, J.M., Ianuzzi, M.C., Kerem, B., Drumm, M.L.,
Melmer, G., Dean, M., Rozmahel, R., Cole, J.L., Kennedy,
D., Hidaka, N., Zsiga, M., Burchwald, M., Riordan, J.R.,
30 Tsui, L.-C. & Collins. F.S. (1989) Science 245, 1059-
1065.

Riordan, J.R., Rommens, J.M., Kerem, B., Alon, N.,
Rozmahel, R., Grzelczack, Z., Zielenski, J., Lok, S.,
35 Plavsic, N., Chou, J.-L., Drumm, M.L., Iannuzzi, M.C.,
Collins, Tsui, L.-C. (1989) Science 245, 1066-1072.

36

Kerem, B., Riommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M. & Tsui, L.-C. (1989) Science 245, 1073-1080).

5   Dean, M., White, M.B., Amos, J., Gerrard, B, Stewart, C., Khaw, K.-T., & Leppert, M. (1990) Cell 61, 863-870.

Cutting, G.R., Kasch, L.M., Rosenstein, B.J., Zielenski, J., Tsui, L.-C., Kazazian, H.H., Jr., & Antonarakas, S.E. 10   (1990) Nature (London) 346, 366-369.

Kerem, B., Zielenski, J., Markiewicz, D., Bozon, D., Gazit, E., Yahav, J., Kennedy, D., Riordan, J.R., Collins, F.S., Rommens, J.M. & Tsui, L.-C. (1990) Proc. 15   Natl. Acad. Sci. 87, 8447-8451.

White, M.B., Amos, J., Hsu, J.M.C., Gerrard, B., Finn, P. & Dean, M. (1990) Nature (London) 344, 665-667.

20   Zielenski, J., Bozon, D., Kerem, B., Markiewicz, D., Durie, P., Rommens, J.M. & Tsui, L.-C. (1991) Genomics 10, 229-235.

Hyde, S.C., Emsley, P., Hartshorn, M.J., Mimmack, M.M., 25   Gileadi, U., Pearce, S.R., Gallagher, M.P., Gill, D.R., Hubbard, R.E. & Higgins, C.F. (1990) Nature (London) 346, 362-365.

Gregory, R.J., Cheng, S.H., Rich, D.P., Marshall, J., 30   Paul, S., Hehir, K., Ostedgaard, L., Klinger, K.W., Welsh, M.J. & Smith, A.E. (1990) Nature (London) 347, 382-386.

Drumm, M.L., Pope, H.A., Clift, W.H., Rommens, J.M., 35   Marvin, S.A., Tsui, L.-C., Collins, F.S., Frizzell, R. & Wilson, J.M. (1990) Cell 62, 1227-1233.

37

Rich, D.P., Anderson, M.P., Gregory, R.J., Cheng, S.H., Paul, S., Jefferson, D.M., McCann, J.D., Klinger, K.W., Smith, A.E. & Welsh, M.J. (1990) Nature (London) 347, 358-363.

5

Kartner, N., Jensen, T.J., Naismith, A.L., Sun, S., Ackerley, C.A., Reyes, E.F., Tsui, L.-C., Rommens, J.M., Bear, C.E., & Riordan, J.R. (1991) Cell.

10   Anderson, M.P., Rich, D.P., Gregrory, R.J., Smith, A.E., Welsh, M.J. (1991) Science 251, 679-682.

Meakin, S.O., Du, R.P., Tsui, L.-C. & Breitman, M.L. (1987) Mol. Cell. Biol. 7, 2671-2679.

15

Karin, M. & Richards, R.I. (1982) Nature (London) 299, 797-802.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989)
20   Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.

Higuchi, R., Krummel, B. & Saiki, R.K. (1988) Nucleic Acids Res. 16, 7351-7363.

25

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. & Pease, L.R. (1989) Gene 77, 51-59.

Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) Nucleic
30   Acids Res. 16, 1215.

MacDonald, R.J., Swift, G.H., Przbyla, A.E. & Chirgwin, J.M. (1987) Methods Enzymol. 152, 219-227.

35   Laemmli, U.K. (1970) Nature (London) 227, 680-685.

38

Towbin, H., Staehelin, T. & Gordon, J. (1979) _Proc. Natl. Acad. Sci. USA_ 76, 4350-4354.

Forskett, J.K. (1988) _Am. J. Physiol._ 255, C566-C571.

Forskett, J.K. & Melvin, J.E. (1989) _Science_ 244, 1582-1585.

Forskett, J.K. (1990) _Am. J. Physiol._ 259, C998-C1004.

Hamill, O.P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F.J. (1981) _Pfügers Arch._ 391, 85-100.

Cheng, S.H., Gregory, R.J., Marshall, J., Paul, S., Souza, D.W., White, G.A., O'Riordan, C.R. & Smith, S.E. (1990) _Cell_ 63, 827-834.

Illsley, N.P. & Verkman, A.S. (1987) _Biochemistry_ 26, 1215-1219.

Cliff, W.H. & Frizell, R.A. (1990) _Proc. Natl Acad. Sci. USA_ 87, 4956-4960.

Tabcharani, J.A., Low, W., Elie, D. & Hanrahan, J.W. (1990) _FEBS Lett._ 270, 157-164.

Gray, M.A., Pollard, C.E., Harris, A., Coleman, L., Greenwell, J.R. & Argent, B.E. (1990) _Am. J. Physiol._ 259, C752-C761.

Ringe, D. & Petsko, G.A. (1990) _Nature (London)_ 346, 312-313.

39

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE
PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

         1.   A modified DNA sequence encoding a cystic fibrosis

5   transmembrane conductance regulator (CFTR) protein, said gene having at
least 27 exons of normal cDNA encoding a CFTR protein of Figure 1, said
normal cDNA including exons 6a and 6b wherein exon 6b includes a 13 bp
repeat,

         said modified DNA sequence comprising at least one of said 13 bp

10   repeats of exon 6b having one or more nucleotide substitutions in said 13 bp
repeat wherein the amino acid sequence of the protein encoded by said DNA
sequence is not changed by said one or more substitutions.

         2.   A modified cDNA sequence of claim 1 wherein said one or more
normal nucleotides are at positions 930 and 933 and said substitutions are C

15   for T at 930 and G for A at 933.

         3.   A DNA construct for use in a recombinant vector comprising
said modified cDNA of claim 1.

         4.   A vector comprising a DNA construct of claim 3, and a promoter
sequence for said DNA construct.

20   5.   A vector of claim 4 wherein said vector comprises a restriction
map of Figure 5 and 8.

         6.   A host cell for producing CFTR protein, said host cell comprising
a compatible vector of claim 4 whereby expression of said vector in said host
cell produces CFTR protein.

25   7.   A host cell of claim 6 selected from the group consisting of
bacterial and mammalian cells.

         8.   A CFTR protein isolated and purified from culture of a host cell
of claim 6.

         9.   A mammalian cell transfected with a vector of claim 4 to

30   enhance C1 conductance through cell wall.

         10.   A modified cDNA sequence of claim 1 further comprising a
nucleotide substitution at position 936, said substitution being C for T.

FIG.1.

```
  1   AATTGGAAGCAAATGACATCACAGCAGGTCAGAGAAAAAGGGTTGAGCGGCAGGCACCCA

 61   GAGTAGTAGGTCTTTGGCATTAGGAGCTTGAGCCCAGACGGCCCTAGCAGGGACCCCAGC

              M  Q  R  S  P  L  E  K  A  S  V  V  S  K  L  F    16
121   GCCCGAGAGACCATGCAGAGGTCGCCTCTGGAAAAGGCCAGCGTTGTCTCCAAACTTTTT

       F  S  W  T  R  P  I  L  R  K  G  Y  R  Q  R  L  E  L  S  D    36
181   TTCAGCTGGACCAGACCAATTTTGAGGAAAGGATACAGACAGCGCCTGGAATTGTCAGAC

       I  Y  Q  I  P  S  V  D  S  A  D  N  L  S  E  K  L  E  R  E    56
241   ATATACCAAATCCCTTCTGTTGATTCTGCTGACAATCTATCTGAAAAATTGGAAAGAGAA

       W  D  R  E  L  A  S  K  K  N  P  K  L  I  N  A  L  R  R  C    76
301   TGGGATAGAGAGCTGGCTTCAAAGAAAAATCCTAAACTCATTAATGCCCTTCGGCGATGT

       F  F  W  R  F  M  F  Y  G  I  F  L  Y  L  G  E  V  T  K  A    96
361   TTTTTCTGGAGATTTATGTTCTATGGAATCTTTTTATATTTAGGGGAAGTCACCAAAGCA

       V  Q  P  L  L  L  G  R  I  I  A  S  Y  D  P  D  N  K  E  E   116
421   GTACAGCCTCTCTTACTGGGAAGAATCATAGCTTCCTATGACCCGGATAACAAGGAGGAA

       R  S  I  A  I  Y  L  G  I  G  L  C  L  L  F  I  V  R  T  L   136
481   CGCTCTATCGCGATTTATCTAGGCATAGGCTTATGCCTTCTCTTTATTGTGAGGACACTG

       L  L  H  P  A  I  F  G  L  H  H  I  G  M  Q  M  R  I  A  M   156
541   CTCCTACACCCAGCCATTTTTGGCCTTCATCACATTGGAATGCAGATGAGAATAGCTATG

       F  S  L  I  Y  K  K  T  L  K  L  S  S  R  V  L  D  K  I  S   176
601   TTTAGTTTGATTTATAAGAAGACTTTAAAGCTGTCAAGCCGTGTTCTAGATAAAATAAGT

       I  G  Q  L  V  S  L  L  S  N  N  L  N  K  F  D  E  G  L  A   196
661   ATTGGACAACTTGTTAGTCTCCTTTCCAACAACCTGAACAAATTTGATGAAGGACTTGCA

       L  A  H  F  V  W  I  A  P  L  Q  V  A  L  L  M  G  L  I  W   216
721   TTGGCACATTTCGTGTGGATCGCTCCTTTGCAAGTGGCACTCCTCATGGGGCTAATCTGG

       E  L  L  Q  A  S  A  F  C  G  L  G  F  L  I  V  L  A  L  F   236
781   GAGTTGTTACAGGCGTCTGCCTTCTGTGGACTTGGTTTCCTGATAGTCCTTGCCCTTTTT

       Q  A  G  L  G  R  M  M  M  K  Y  R  D  Q  R  A  G  K  I  S   256
841   CAGGCTGGGCTAGGGAGAATGATGATGAAGTACAGAGATCAGAGAGCTGGGAAGATCAGT

       E  R  L  V  I  T  S  E  M  I  E  N  I  Q  S  V  K  A  Y  C   276
901   GAAAGACTTGTGATTACCTCAGAAATGATTGAAAATATCCAATCTGTTAAGGCATACTGC

       W  E  E  A  M  E  K  M  I  E  N  L  R  Q  T  E  L  K  L  T   296
961   TGGGAAGAAGCAATGGAAAAAATGATTGAAAACTTAAGACAAACAGAACTGAAACTGACT

       R  K  A  A  Y  V  R  Y  F  N  S  S  A  F  F  F  S  G  F  F   316
1021  CGGAAGGCAGCCTATGTGAGATACTTCAATAGCTCAGCCTTCTTCTTCTCAGGGTTCTTT

       V  V  F  L  S  V  L  P  Y  A  L  I  K  G  I  I  L  R  K  I   336
1081  GTGGTGTTTTTATCTGTGCTTCCCTATGCACTAATCAAAGGAATCATCCTCCGGAAAATA

       F  T  T  I  S  F  C  I  V  L  R  M  A  V  T  R  Q  F  P  W   356
1141  TTCACCACCATCTCATTCTGCATTGTTCTGCGCATGGCGGTCACTCGGCAATTTCCCTGG

       A  V  Q  T  W  Y  D  S  L  G  A  I  N  K  I  Q  D  F  L  Q   376
1201  GCTGTACAAACATGGTATGACTCTCTTGGAGCAATAAACAAAATACAGGATTTCTTACAA

       K  Q  E  Y  K  T  L  E  Y  N  L  T  T  T  E  V  V  M  E  N   396
1261  AAGCAAGAATATAAGACATTGGAATATAACTTAACGACTACAGAAGTAGTGATGGAGAAT

       V  T  A  F  W  E  E  G  F  G  E  L  F  E  K  A  K  Q  N  N   416
1321  GTAACAGCCTTCTGGGAGGAGGGATTTGGGGAATTATTTGAGAAAGCAAAACAAAACAAT
```

FIG.1(cont'd)

```
              N  N  R  K  T  S°  N  G  D  D  S  L  F  F  S  N  F  S  L  L    436
1381  AACAATAGAAAAACTTCTAATGGTGATGACAGCCTCTTCTTCAGTAATTTCTCACTTCTT

              G  T  P  V  L  K  D  I  N  F  K  I  E  R  G  Q  L  L  A  V    456
1441  GGTACTCCTGTCCTGAAAGATATTAATTTCAAGATAGAAAGAGGACAGTTGTTGGCGGTT

              A  G  S  T  G  A  G  K| T  S  L  L  M  H  I  M  G  E  L  E    476
1501  GCTGGATCCACTGGAGCAGGCAAGACTTCACTTCTAATGATGATTATGGGAGAACTGGAG

              P  S  E  G  K  I  K  H  S°  G  R  I  S  F  C  S  Q  F  S  W    496
1561  CCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTCATTCTGTTCTCAGTTTTCCTGG

              I  M  P  G  T° I  K  E  N  I  I  F△  G  V  S  Y  D  E  Y  R    516
1621  ATTATGCCTGGCACCATTAAAGAAAATATCATCTTTGGTGTTTCCTATGATGAATATAGA

              Y  R  S  V  I  K  A  C  Q  L  E  E| D  I  S  K  F  A  E  K    536
1681  TACAGAAGCGTCATCAAAGCATGCCAACTAGAAGAGGACATCTCCAAGTTTGCAGAGAAA

              D  N  I  V  L  G  E  G  G  I  T  L  S  G  G  Q  R  A  R  I    556
1741  GACAATATAGTTCTTGGAGAAGGTGGAATCACACTGAGTGGAGGTCAACGAGCAAGAATT

              S  L  A  R| A  V  Y  K  D  A  D  L  Y  L  L  D  S  P  F  G    576
1801  TCTTTAGCAAGAGCAGTATACAAAGATGCTGATTTGTATTTATTAGACTCTCCTTTTGGA

              Y  L  D  V  L  T° E  K  E  I  F  E  S  C  V  C  K  L  M  A    596
1861  TACCTAGATGTTTTAACAGAAAAAGAAATATTTGAAGCTGTGTCTGTAAACTGATGGCT

              N  K  T  R  I  L  V  T  S  K  M  E  H  L  K  K  A  D  K  I    616
1921  AACAAAACTAGGATTTTGGTCACTTCTAAAATGGAACATTTAAAGAAAGCTGACAAAATA

              L  I  L  H  E  G  S° S  Y  F  Y  G  T  F  S  E  L  Q  N  L    636
1981  TTAATTTTGCATGAAGGTAGCAGCTATTTTTATGGGACATTTTCAGAACTCCAAAATCTA

              Q  P  D  F  S  S° K  L  M  G  C  D  S  F  D  Q  F  S  A  E    656
2041  CAGCCAGACTTTAGCTCAAAACTCATGGGATGTGATTCTTTCGACCAATTTAGTGCAGAA

              R  R  N  S  I  L  T  E  T  L  H  R  F  S  L  E  G  D  A  P    676
2101  AGAAGAAATTCAATCCTAACTGAGACCTTACACCGTTTCTCATTAGAAGGAGATGCTCCT

              V  S  W  T  E  T° K  K  Q  S°°  F  K  Q  T  G  E  F  G  E  K    696
2161  GTCTCCTGGACAGAAACAAAAAAACAATCTTTTAAACAGACTGGAGAGTTTGGGGAAAAA

              R  K  N  S° I  L  N  P  I  N  S° I  R  K  F  S° I  V  Q  K    716
2221  AGGAAGAATTCTATTCTCAATCCAATCAACTCTATACGAAAATTTTCCATTGTGCAAAAG

              T  P  L  Q  M  N  G  I  E  E  D  S  D  E  P  L  E  R  R  L    736
2281  ACTCCCTTACAAATGAATGGCATCGAAGAGGATTCTGATGAGCCTTTAGAGAGAAGGCTG

              S° L  V  P  D  S  E  Q  G  E  A  I  L  P  R  I  S  V  I  S    756
2341  TCCTTAGTACCAGATTCTGAGCAGGGAGAGGCGATACTGCCTCGCATCAGCGTGATCAGC

              T  G  P  T  L  Q  A  R  R  R  Q  S° V  L  N  L  M  T  H  S    776
2401  ACTGGCCCCACGCTTCAGGCACGAAGGAGGCAGTCTGTCCTGAACCTGATGACACACTCA

              V  N  Q  G  Q  N  I  H  R  K  T  T° A  S° T° R  K  V  S° L    796
2461  GTTAACCAAGGTCAGAACATTCACCGAAAGACAACAGCATCCACACGAAAAGTGTCACTG

              A  P  Q  A  N  L  T  E  L  D  I  Y  S° R  R  L  S° Q  E  T    816
2521  GCCCCTCAGGCAAACTTGACTGAACTGGATATATATTCAAGAAGGTTATCTCAAGAAACT
```

A

```
               G  L  E  I  S  E  E  I  N  E  E  D  L  K │ E  C  F  F  D  D    836
     2581  GGCTTGGAAATAAGTGAAGAAATTAACGAAGAAGACTTAAAGGAGTGCTTTTTTGATGAT

               M  E  S  I  P  A  V  T  T  W  N  T  Y  L  R  Y  I  T  V  H    856
     2641  ATGGAGAGCATACCAGCAGTGACTACATGGAACACATACCTTCGATATATTACTGTCCAC

               K  S  L  I  F  V  L  I  W  C  L  V  I  F  L  A  E  V  A  A    876
     2701  AAGAGCTTAATTTTTGTGCTAATTTGGTGCTTAGTAATTTTTCTGGCAGAGGTGGCTGCT

               S  L  V  V  L  W  L  L  G  M  T  P  L  Q  D  K  G  N  S  T    896
     2761  TCTTTGGTTGTGCTGTGGCTCCTTGGAAACACTCCTCTTCAAGACAAAGGGAATAGTACT

               H  S  R  N  N  S  Y  A  V  I  I  T  S  T  S  S  Y  Y  V  F    916
     2821  CATAGTAGAAATAACAGCTATGCAGTGATTATCACCAGCACCAGTTCGTATTATGTGTTT

               Y  I  Y  V  G  V  A  D  T  L  L  A  M  G  F  F  R  G  L  P    936
     2881  TACATTTACGTGGGAGTAGCCGACACTTTGCTTGCTATGGGATTCTTCAGAGGTCTACCA

               L  V  H  T  L  I  T  V  S  K  I  L  H  H  K  M  L  H  S  V    956
     2941  CTGGTGCATACTCTAATCACAGTGTCGAAAATTTTACACCACAAAATGTTACATTCTGTT

               L  Q  A  P  M  S  T  L  N  T  L  K  A  G  I  L  N  R  F     976
     3001  CTTCAAGCACCTATGTCAACCCTCAACACGTTGAAAGCAGGTGGGATTCTTAATAGATTC

               S  K  D  I  A  I  L  D  D  L  L  P  L  T  I  F  D  F  I  Q    996
     3061  TCCAAAGATATAGCAATTTTGGATGACCTTCTGCCTCTTACCATATTTGACTTCATCCAG

               L  L  L  I  V  I  G  A  I  A  V  V  A  V  L  Q  P  Y  I  F   1016
     3121  TTGTTATTAATTGTGATTGGAGCTATAGCAGTTGTCGCAGTTTTACAACCCTACATCTTT

               V  A  T  V  P  V  I  V  A  F  I  M  L  R  A  Y  F  L  Q  T   1036
     3181  GTTGCAACAGTGCCAGTGATAGTGGCTTTTATTATGTTGAGAGCATATTTCCTCCAAACC

               S  Q  Q  L  K  Q  L  E  S  E  G  R  S  P  I  F  T  H  L  V   1056
     3241  TCACAGCAACTCAAACAACTGGAATCTGAAGGCAGGAGTCCAATTTTCACTCATCTTGTT

               T  S  L  K  G  L  W  T  L  R  A  F  G  R  Q  P  Y  F  E  T   1076
     3301  ACAAGCTTAAAAGGACTATGGACACTTCGTGCCTTCGGACGGCAGCCTTACTTTGAAACT

               L  F  H  K  A  L  N  L  H  T  A  N  W  F  L  Y  L  S  T  L   1096
     3361  CTGTTCCACAAAGCTCTGAATTTACATACTGCCAACTGGTTCTTGTACCTGTCAACACTG

               R  W  F  Q  M  R  I  E  M  I  F  V  I  F  F  I  A  V  T  F   1116
     3421  CGCTGGTTCCAAATGAGAATAGAAATGATTTTTGTCATCTTCTTCATTGCTGTTACCTTC

               I  S  I  L  T  T  G  E  G  E  G  R  V  G  I  I  L  T  L  A   1136
     3481  ATTTCCATTTTAACAACAGGAGAAGGAGAAGGAAGAGTTGGTATTATCCTGACTTTAGCC

               M  N  I  M  S  T  L  Q  W  A  V  N  S  S  I  D  V  D  S  L   1156
     3541  ATGAATATCATGAGTACATTGCAGTGGGCTGTAAACTCCAGCATAGATGTGGATAGCTTG

               M  R  S  V  S  R  V  F  K  F  I  D  M  P  T  E  G  K  P  T   1176
     3601  ATGCGATCTGTGAGCCGAGTCTTTAAGTTCATTGACATGCCAACAGAAGGTAAACCTACC

               K  S  T  K  P  Y  K  N  G  Q  L  S  K  V  M  I  I  E  N  S   1196
     3661  AAGTCAACCAAACCATACAAGAATGGCCAACTCTCGAAAGTTATGATTATTGAGAATTCA

               H  V  K  K  D  D  I  W  ·P  S  G  G  Q  M  T  V  K  D  L  T   1216
     3721  CACGTGAAGAAAGATGACATCTGGCCCTCAGGGGGCCAAATGACTGTCAAAGATCTCACA

               A  K  Y  T  E  G  G  N  A  I  L  E  N  I  S  F  S  I  S  P   1236
     3781  GCAAAATACACAGAAGGTGGAAATGCCATATTAGAGAACATTTCCTTCTCAATAAGTCCT

               G  Q  R  V  G  L  L  G  R  T  G  S  G  K  S  T  L  L  S  A   1256
     3841  GGCCAGAGGGTGGGCCTCTTGGGAAGAACTGGATCAGGGAAGAGTACTTTGTTATCAGCT
```

FIG.1. (cont'd)

```
          F   L   R   L   L   N   T   E   G   E   I   Q   I   D   G   V   S   W   D   S   1276
3901  TTTTTGAGACTACTGAACACTGAAGGAGAAATCCAGATCGATGGTGTGTCTTGGGATTCA

          I   T   L   Q   Q   W   R   K   A   F   G   V   I   P   Q   K   V   F   I   F   1296
3961  ATAACTTTGCAACAGTGGAGGAAAGCCTTTGGAGTGATACCACAGAAAGTATTTATTTTT

          S   G   T   F   R   K   N   L   D   P   Y   E   Q   W   S   D   Q   E   I   W   1316
4021  TCTGGAACATTTAGAAAAAACTTGGATCCCTATGAACAGTGGAGTGATCAAGAAATATGG

          K   V   A   D   E   V   G   L   R   S   V   I   E   Q   F   P   G   K   L   D   1336
4081  AAAGTTGCAGATGAGGTTGGGCTCAGATCTGTGATAGAACAGTTTCCTGGGAAGCTTGAC

          F   V   L   V   D   G   G   C   V   L   S   H   G   H   K   Q   L   M   C   L   1356
4141  TTTGTCCTTGTGGATGGGGGCTGTGTCCTAAGCCATGGCCACAAGCAGTTGATGTGCTTG

          A   R   S   V   L   S   K   A   K   I   L   L   L   D   E   P   S   A   H   L   1376
4201  GCTAGATCTGTTCTCAGTAAGGCGAAGATCTTGCTGCTTGATGAACCCAGTGCTCATTTG

          D   P   V   T   Y   Q   I   I   R   R   T   L   K   Q   A   F   A   D   C   T   1396
4261  GATCCAGTAACATACCAAATAATTAGAAGAACTCTAAAACAAGCATTTGCTGATTGCACA

          V   I   L   C   E   H   R   I   E   A   M   L   E   C   Q   Q   F   L   V   I   1416
4321  GTAATTCTCTGTGAACACAGGATAGAAGCAATGCTGGAATGCCAACAATTTTTGGTCATA

          E   E   N   K   V   R   Q   Y   D   S   I   Q   K   L   L   N   E   R   S   L   1436
4381  GAAGAGAACAAAGTGCGGCAGTACGATTCCATCCAGAAACTGCTGAACGAGAGGAGCCTC

          F   R   Q   A   I   S   P   S   D   R   V   K   L   F   P   H   R   N   S   S   1456
4441  TTCCGGCAAGCCATCAGCCCCTCCGACAGGGTGAAGCTCTTTCCCCACCGGAACTCAAGC

          K   C   K   S   K   P   Q   I   A   A   L   K   E   E   T   E   E   E   V   Q   1476
4501  AAGTGCAAGTCTAAGCCCCAGATTGCTGCTCTGAAAGAGGAGACAGAAGAAGAGGTGCAA

          D   T   R   L   =                                                             1480
4561  GATACAAGGCTTTAGAGAGCAGCATAAATGTTGACATGGGACATTTGCTCATGGAATTGG
4621  AGCTCGTGGGACAGTCACCTCATGGAATTGGAGCTCGTGGAACAGTTACCTCTGCCTCAG
4681  AAAACAAGGATGAATTAAGTTTTTTTTTAAAAAAGAAACATTTGGTAAGGGGAATTGAGG
4741  ACACTGATATGGGTCTTGATAAATGGCTTCCTGGCAATAGTCAAATTGTGTGAAAGGTAC
4801  TTCAAATCCTTGAAGATTTACCACTTGTGTTTTGCAAGCCAGATTTTCCTGAAAACCCTT
4861  GCCATGTGCTAGTAATTGGAAAGGCACCTCTAAATGTCAATCAGCCTAGTTGATCAGCTT
4921  ATTGTCTAGTGAAACTCGTTAATTTGTAGTGTTGGAGAAGAACTGAAATCATACTTCTTA
4981  GGGTTATGATTAAGTAATGATAACTGGAAACTTCAGCGGTTTATATAAGCTTGTATTCCT
5041  TTTTCTCTCCTCTCCCCATGATGTTTAGAAACACAACTATATTGTTTGCTAAGCATTCCA
5101  ACTATCTCATTTCCAAGCAAGTATTAGAATACCACAGGAACCACAAGACTGCACATCAAA
5161  ATATGCCCCATTCAACATCTAGTGAGCAGTCAGGAAAGAGAACTTCCAGATCCTGGAAAT
5221  CAGGGTTAGTATTGTCCAGGTCTACCAAAAATCTCAATATTTCAGATAATCACAATACAT
5281  CCCTTACCTGGGAAAGGGCTGTTATAATCTTTCACAGGGGACAGGATGGTTCCCTTGATG
5341  AAGAAGTTGATATGCCTTTTCCCAACTCCAGAAAGTGACAAGCTCACAGACCTTTGAACT
5401  AGAGTTTAGCTGGAAAAGTATGTTAGTGCAAATTGTCACAGGACAGCCCTTCTTTCCACA
5461  GAAGCTCCAGGTAGAGGGTGTGTAAGTAGATAGGCCATGGGCACTGTGGGTAGACACACA
5521  TGAAGTCCAAGCATTTAGATGTATAGGTTGATGGTGGTATGTTTTCAGGCTAGATGTATG
5581  TACTTCATGCTGTCTACACTAAGAGAGAATGAGAGACACACTGAAGAAGCACCAATCATG
5641  AATTAGTTTTATATGCTTCTGTTTTATAATTTTGTGAAGCAAAATTTTTTCTCTAGGAAA
5701  TATTTATTTTAATAATGTTTCAAACATATATTACAATGCTGTATTTAAAAGAATGATTA
5761  TGAATTACATTTGTATAAAATAATTTTTATATTTGAAATATTGACTTTTTATGGCACTAG
5821  TATTTTTATGAAATATTATGTTAAAACTGGGACAGGGGAGAACCTAGGGTGATATTAACC
5881  AGGGGCCATGAATCACCTTTTGGTCTGGAGGGAAGCCTTGGGGCTGATCGAGTTGTTGCC
5941  CACAGCTGTATGATTCCCAGCCAGACACAGCCTCTTAGATGCAGTTCTGAAGAAGATGGT
6001  ACCACCAGTCTGACTGTTTCCATCAAGGGTACACTGCCTTCTCAACTCCAAACTGACTCT
6061  TAAGAAGACTGCATTATATTTATTACTGTAAGAAAATATCACTTGTCAATAAAATCCATA
6121  CATTTGTGT (A) n
```
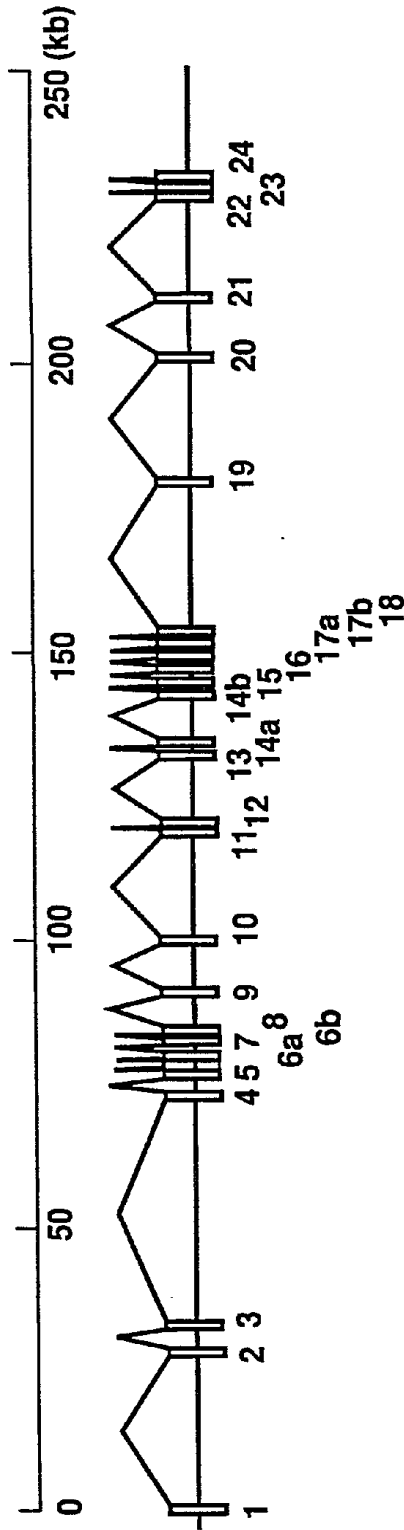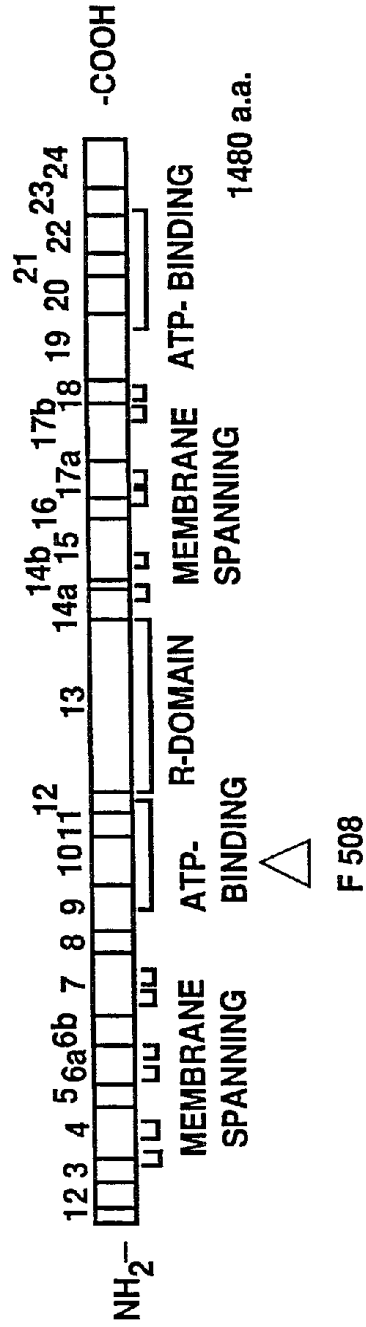
FIG. 2a. CF GENE

FIG. 2b. CFTR

FIG.3.

Exon 6b:

```
        876
        AGATCAGAGAGCTGGGAAGATCAGTGAAAGACTTGTGATTACCTCAG

        -35                                    -10
        TTGACA--------(17±1 bp)--------TATAAT
        AAATGATTGAAAATATCCAATCTGTTAAGGCATACTGCTGGGAAGAAGCAATGGAAA
                   C  G  C
                  930 933 936

                                                        1027
        AAATGATTGAAAACTTAAGACAAACAGAACTGAAACTGACTCGGAAGG
```
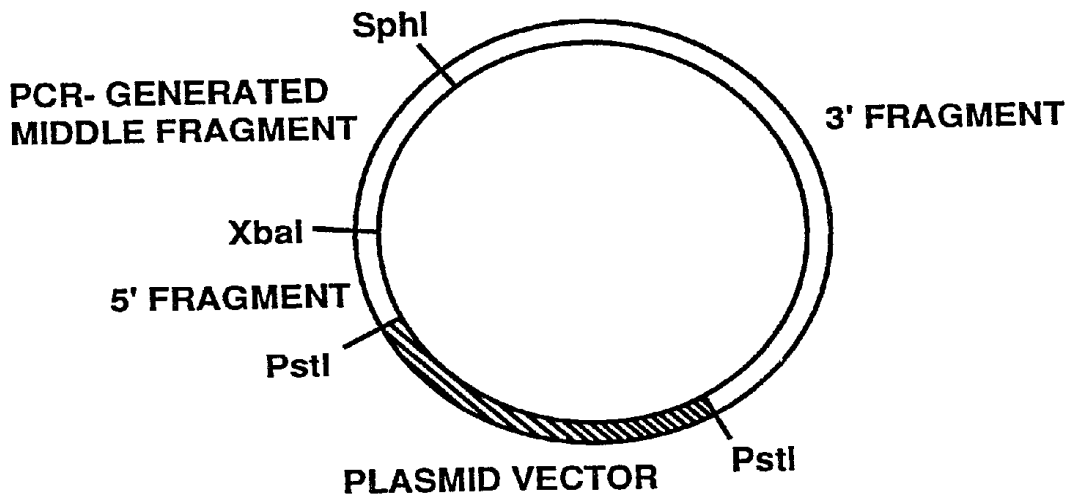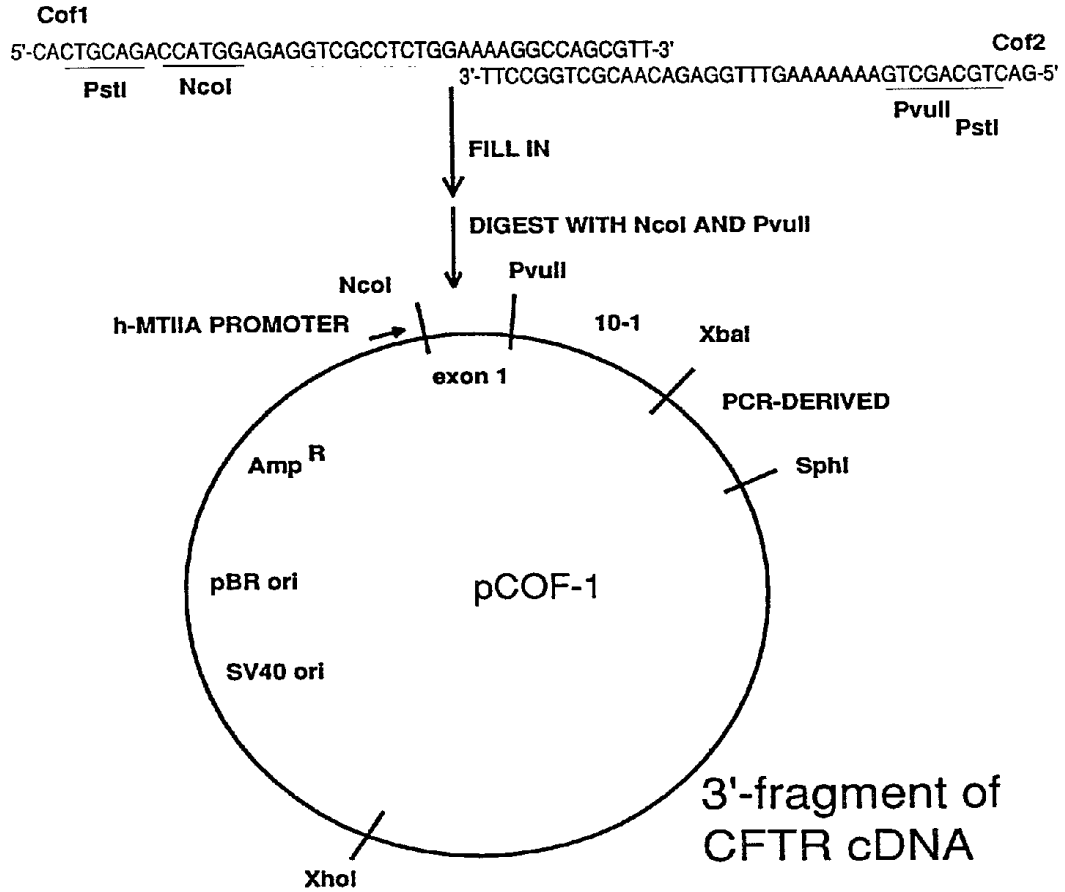
7/20

FIG. 4.

8/20

Cof1
5'-CACTGCAGACCATGGAGAGGTCGCCTCTGGAAAAGGCCAGCGTT-3'   Cof2
                                           3'-TTCCGGTCGCAACAGAGGTTTGAAAAAAAGTCGACGTCAG-5'
   Pstl   Ncol   Pvull
                                                                           Pstl

FILL IN

DIGEST WITH Ncol AND Pvull

Pvull

Ncol

h-MTIIA PROMOTER   10-1   Xbal

exon 1   PCR-DERIVED

                                 Sphl

Amp R

pBR ori   pCOF-1

SV40 ori

                    3'-fragment of
                    CFTR cDNA
Xhol

FIG. 5.

2091907

FROM TE27

Ncol   EcoRI

FROM T16-4.5

BstXI   Ncol

SphI   FROM C1-1/5   BstXI   BstXI   EcoRI
                                                                  PstI

SphI   EcoRI

FIG. 6.

10/20

5'-GCTACCAAAGCAGTACAGCCT-3'

3'-GAGTCTTTACTAGCTCTTATAGGTTAGA-5'

+

5'-CAGAAATGATCGAGAATATCCAATCT-3'

3'-CACCTTAGTGTGACTCACCTGTAG-5'

Xbal   Sphl
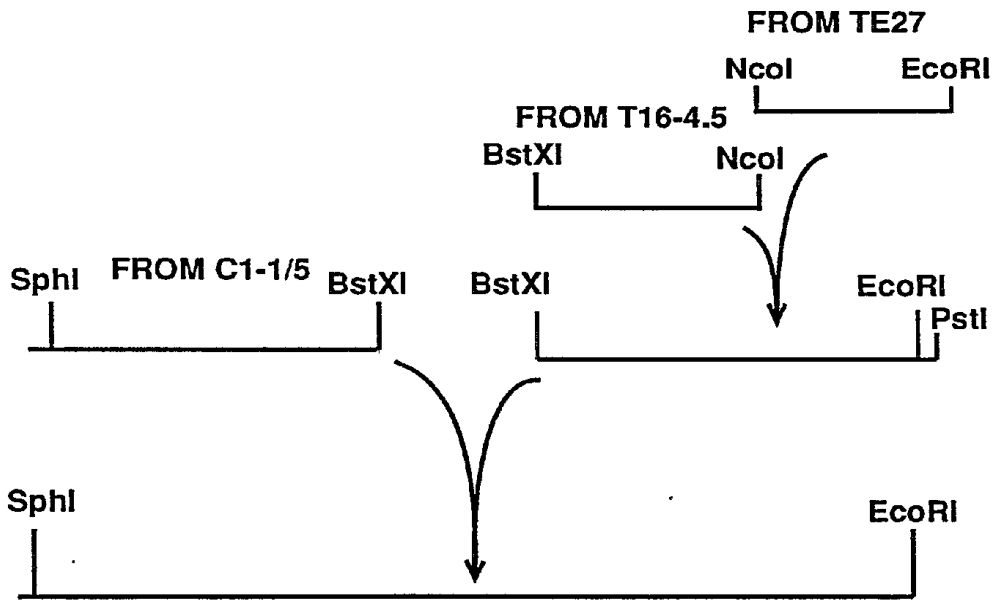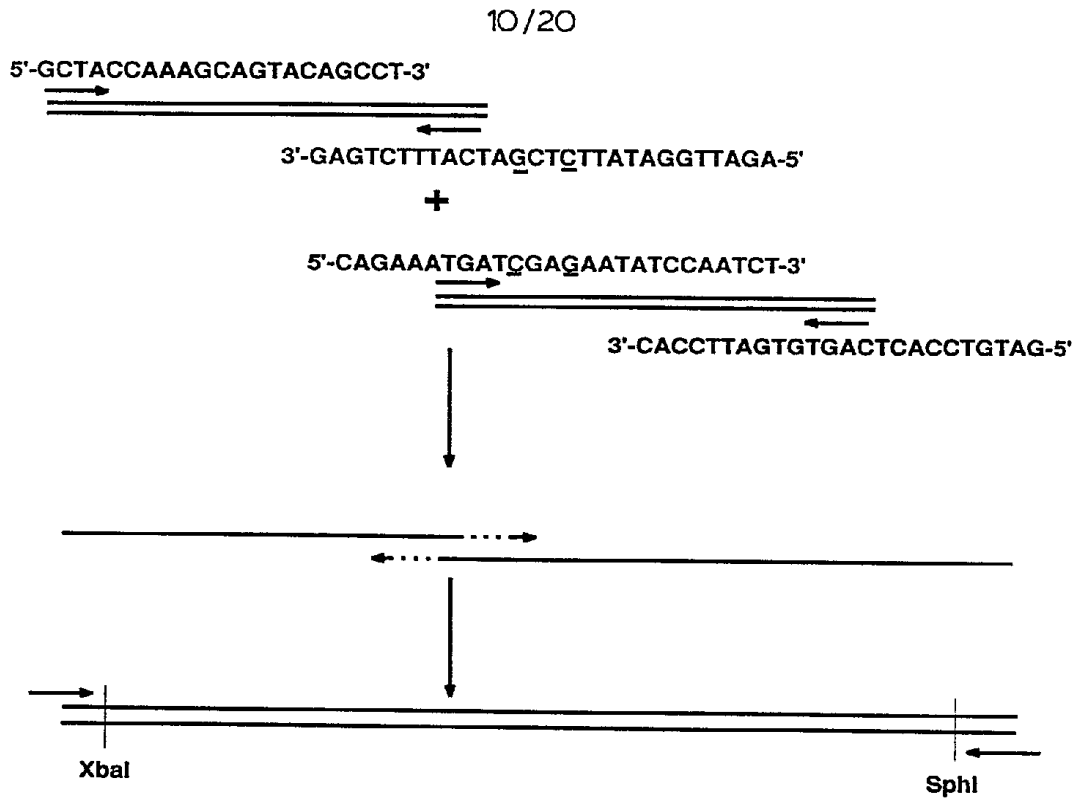
FIG.7.
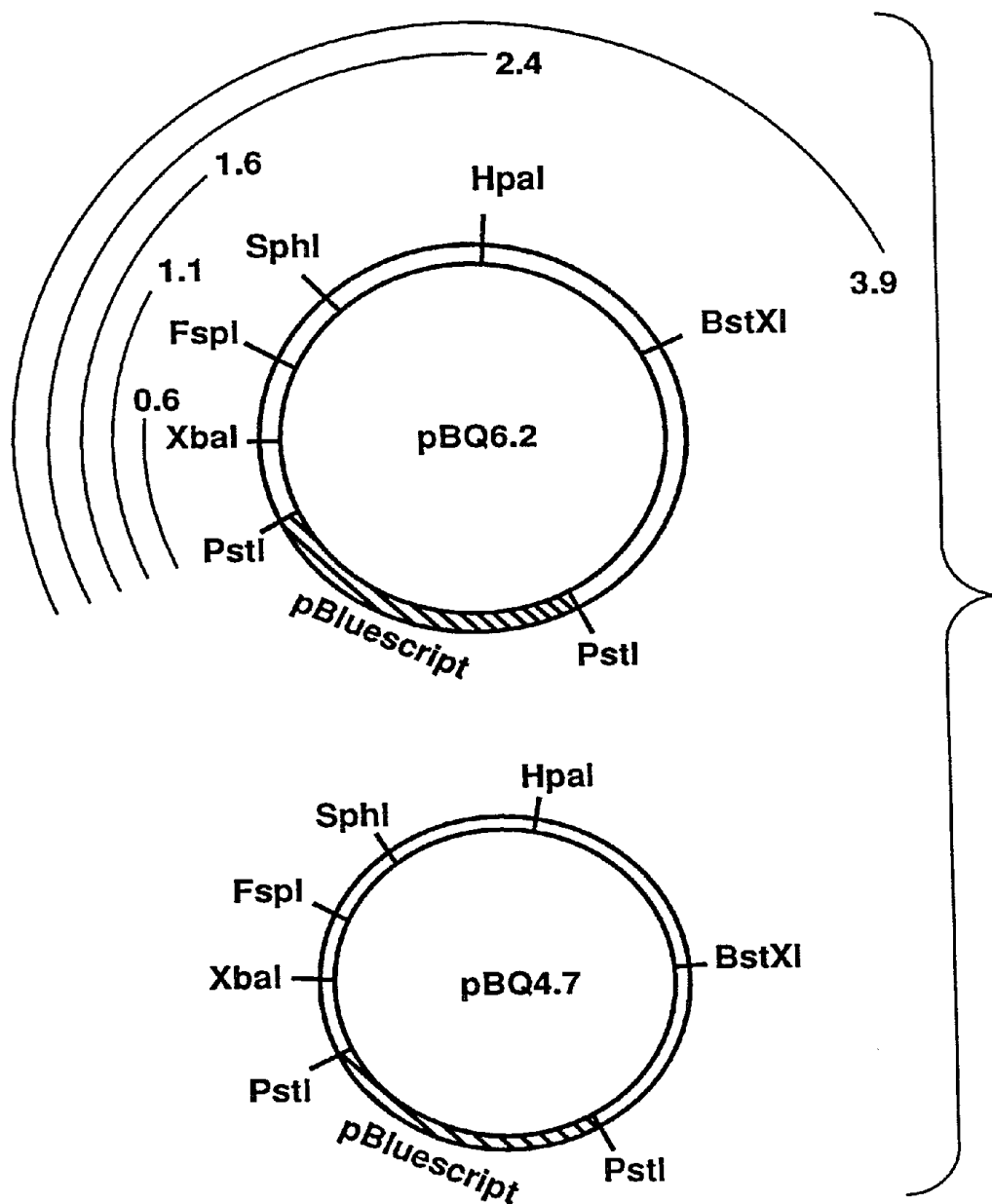
FIG. 8.   SUBSTITUTE SHEET

FIG.10.

FIG.9.
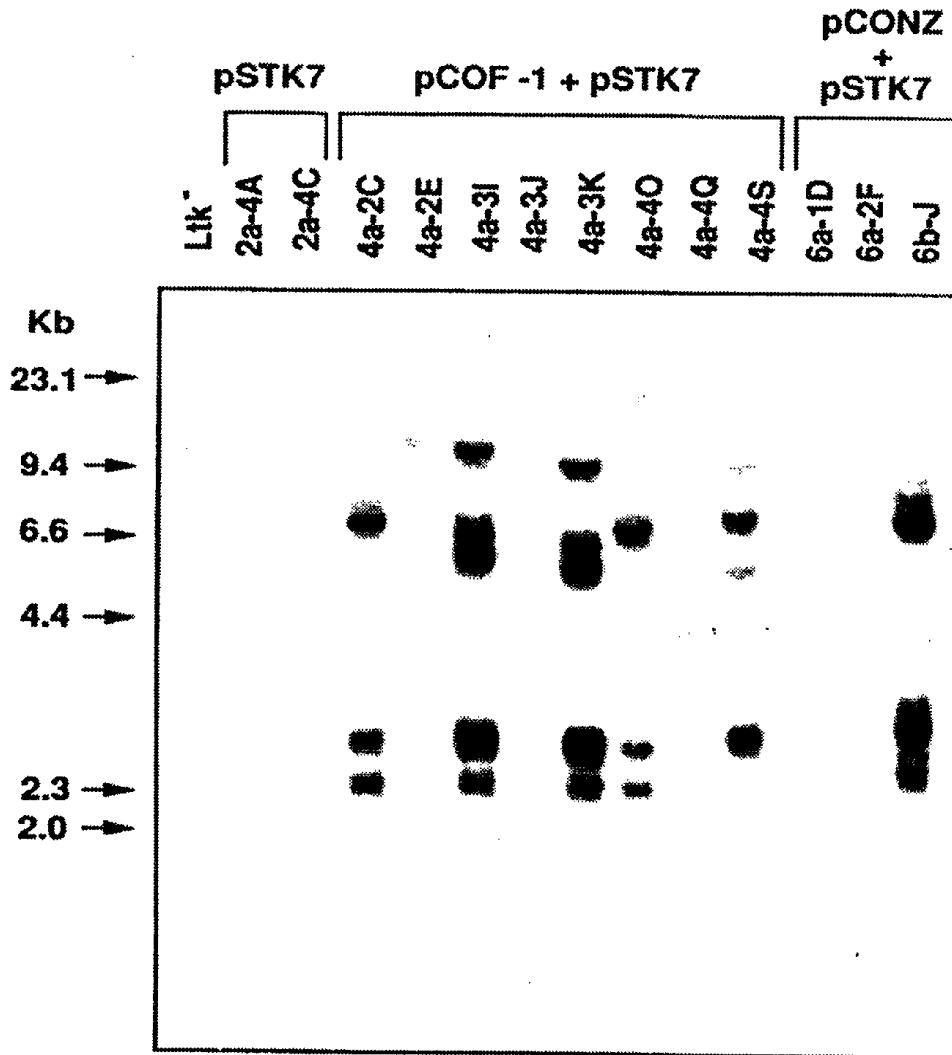
Sequence of Exon 6b in Constructed Vectors

110
97
86.2
35

1 2 3 4

200
97.4
68
43
29

894

GATC

CTGATC
GTCTTTCA
AATCACAA
TCTGAGGT
CGATCATT
GATGTTCT
ACAGATTG
ATGCCTTA
CCAGCAGT
GCTTCTTC
TTTCCATT

979

981

GATC

TGGAAAAA
AGAAGCAA
TGCTGGGA
AGGCATAC
ATCTGTTA
AACATCCA
TGATGGAG
CTCAGAAA
GTGATTAC
AAAGACTT
GATCAGTG

894

DNA ANALYSIS

FIG.11.

FIGURE 14.

Membrane   Whole cell

PANC1 T84 6b-J 4a-3I 4a-2D 2a-4A

— 200

←

FIG.13.

— 97.4
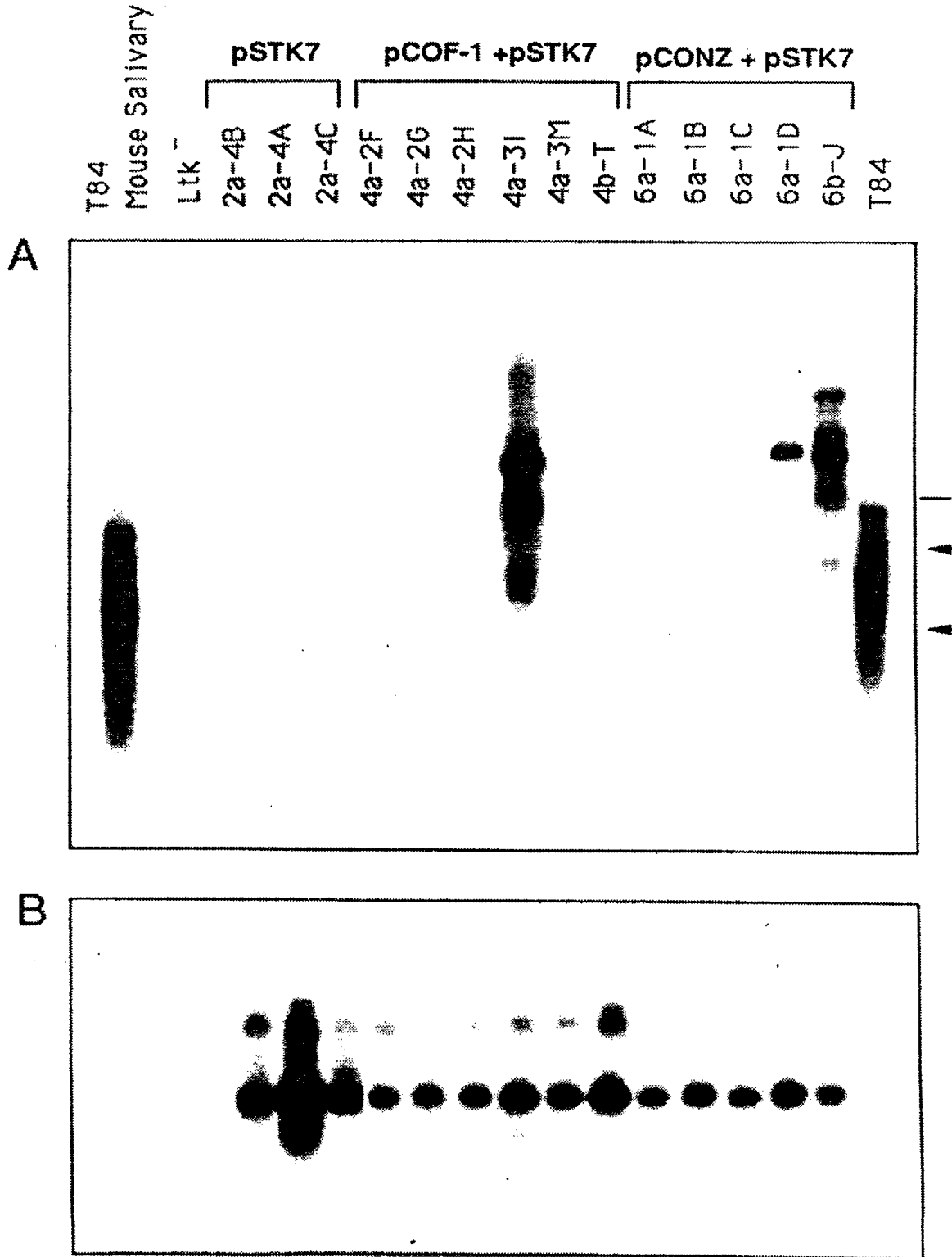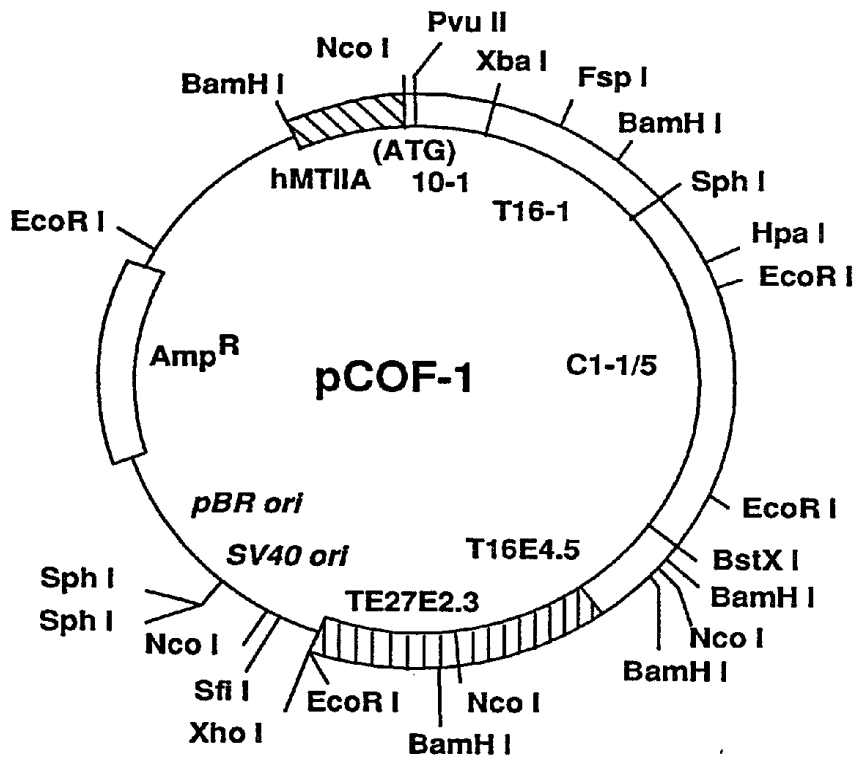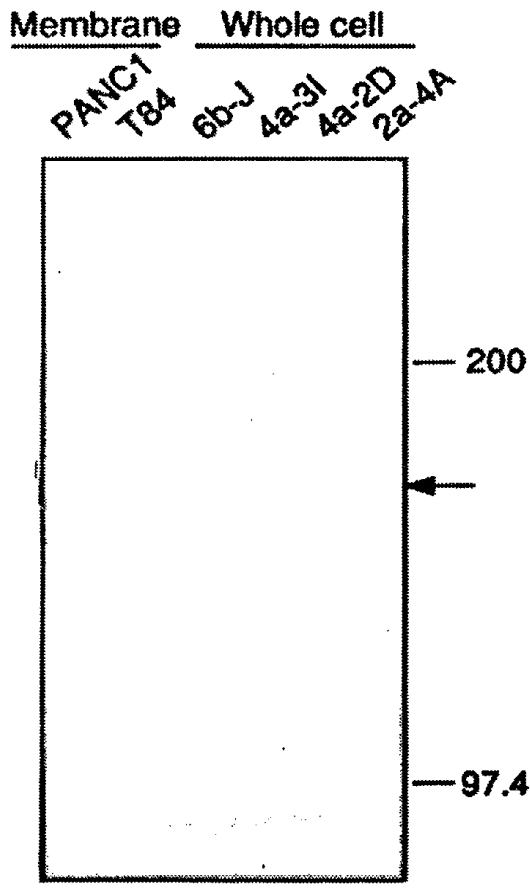
FIG.15.

CaCo2 LTK⁻ 6bI 4a3I 5-2D

←

—

—

T84 LTK⁻   6bI   4a3I   5-2D
A B   A B   A B   A B

200 --

116 --
97 --

66 --

17/20

TRUNCATED CFTR

FIG.16A.

ΔF508

FIG.16B.

CFTR
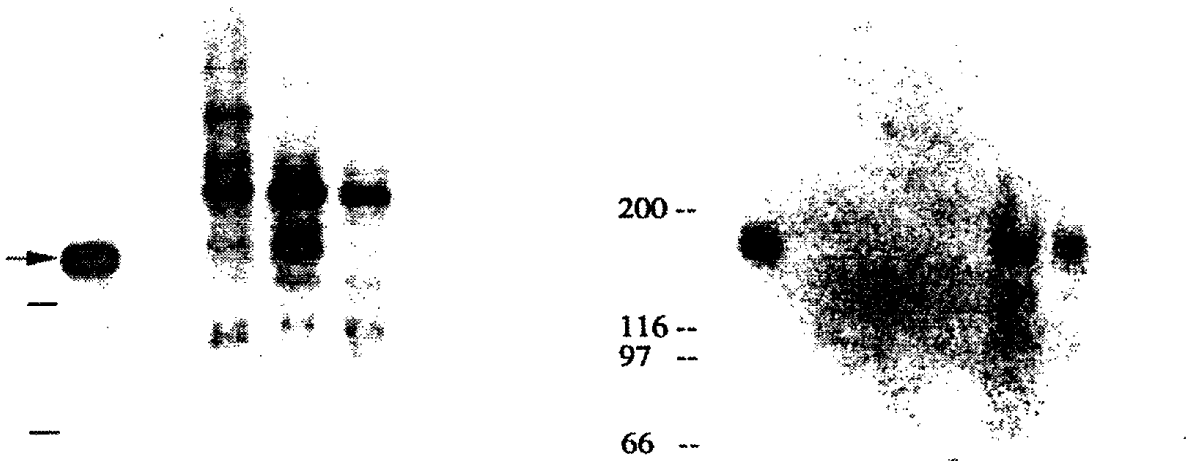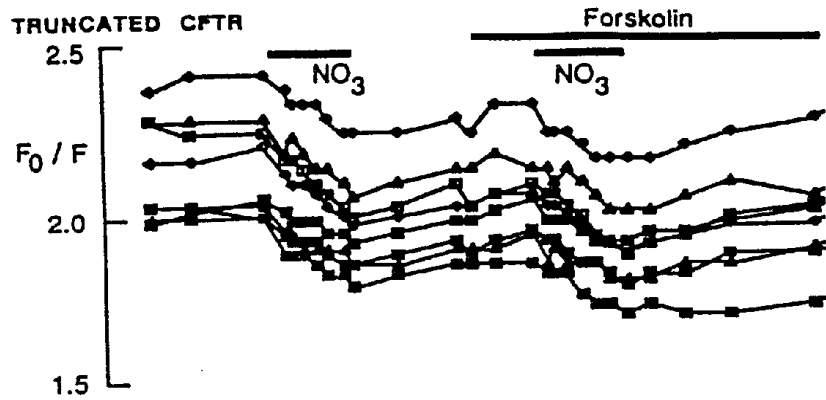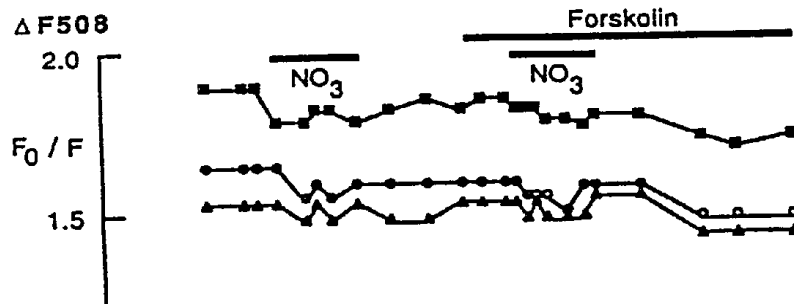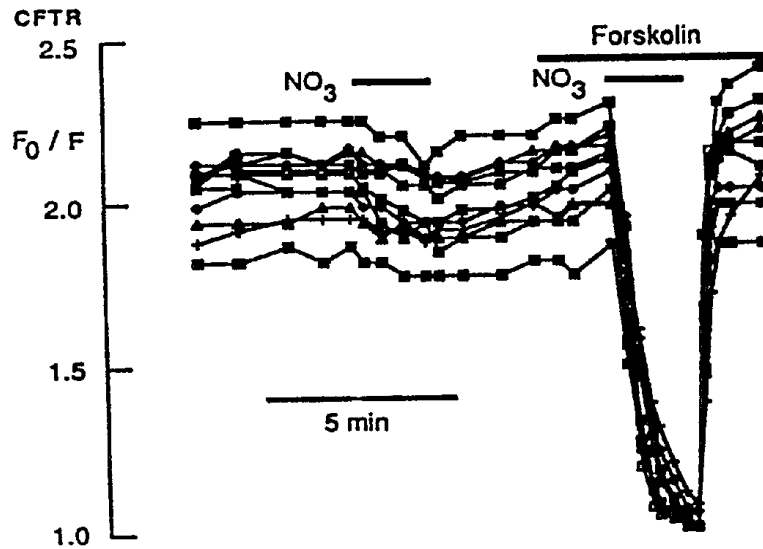
FIG.16C.

2091907

18/20

FIG.16D.

FIG.16E.

FIG. 17A.

FIG.17B.

2091907

20/20

FIG. 17C.