A Novel Exon in the Cystic Fibrosis Transmembrane Conductance Regulator Gene Activated by the Nonsense Mutation E92X in Airway Epithelial Cells of Patients with Cystic Fibrosis

Katrin Will,* Thilo Dörk,‡ Manfred Stuhrmann,* Thomas Meitinger,§ Rosemarie Bertele-Harms,‖ Burkhard Tümmler,‡ and Jörg Schmidtke*

**Abteilung für Humangenetik and ‡Klinische Forschergruppe, Molekulare Pathologie der Mukoviszidose, Medizinische Hochschule, 30623 Hannover; and §Abteilung für pädiatrische Genetik and ‖von Haunersches Kinderspital, Ludwig-Maximilians-Universität, 80336 München, Germany*

Abstract

Cystic fibrosis (CF) is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. We report on a novel nonsense mutation that leads to exon skipping and the activation of a cryptic exon. Screening of genomic DNA from 700 German patients with CF uncovered four cases with the nonsense mutation E92X, a G → T transversion that creates a termination codon and affects the first base of exon 4 of the CFTR gene. Lymphocyte RNA of two CF patients heterozygous for E92X was found to contain the wild type sequence and a differentially spliced isoform lacking exon 4. In RNA derived from nasal epithelial cells of E92X patients, a third fragment of longer size was observed. Sequencing revealed the presence of E92X and an additional 183-bp fragment, inserted between exons 3 and 4. The 183-bp sequence was mapped to intron 3 of the CFTR gene. It is flanked by acceptor and donor splice sites. We conclude that the 183-bp fragment in intron 3 is a cryptic CFTR exon that can be activated in epithelial cells by the presence of the E92X mutation. E92X abolishes correctly spliced CFTR mRNA and leads to severe cystic fibrosis. (*J. Clin. Invest.* 1994. 93:1852–1859.) Key words: cryptic exon • splice site mutation • exon skipping • CFTR • vectorette PCR

Introduction

Cystic fibrosis (CF)[1] is the most common severe autosomal recessive disorder in Caucasian populations. In CF patients, altered chloride transport across the apical membrane of epithelial cells is associated with the accumulation of abnormal mucus in lung and pancreas resulting in chronic obstructive lung disease and pancreatic insufficiency (1, 2). The gene affected in CF, the cystic fibrosis transmembrane conductance regulator (CFTR), spans ~ 250 kb on chromosome 7q31 and comprises 27 exons (3, 4), encoding a cAMP- and ATP-regulated low conductance chloride channel (5). In addition to the major mutation, ΔF508 (6), > 300 further mutations within the CFTR gene have been identified (Cystic Fibrosis Genetic Analysis Consortium, August 1993). These mutations have been shown to cause mRNA deficiency (7–11), aberrant splicing (12–14), defective maturation of the protein (15, 16), and the alteration or loss of chloride channel function (17, 18). In this study, we have identified and characterized a novel nonsense mutation, E92X, that affects the first nucleotide of exon 4, thereby leading to exon skipping and the insertion of a previously unknown cryptic exon into epithelial CFTR mRNA. Case reports of four German patients and the result of transcript analyses indicate that the E92X nonsense mutation causes severe CF because it abolishes correct splicing of the CFTR mRNA.

Methods

Samples. Peripheral blood samples were taken from 700 CF patients of German origin after informed consent had been obtained, and genomic DNA was extracted by standard procedures. Each DNA sample was routinely screened for the presence of the most common mutations as previously described (19, 20). Samples with at least one unidentified mutation were subjected to single strand conformation polymorphism (SSCP) analysis. Human pancreas total RNA was obtained from Clontech (Palo Alto, CA). Colon cell lines T84 and HT-29 were cultured as described previously (21).

SSCP Analysis (22). Approximately 0.5 μg of genomic DNA was included as the template in a 50-μl PCR containing 20 μM dNTPs, 0.5 μM of each primer, 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris/HCl, pH 8.0, and 1.5 U Taq DNA polymerase. After 20 PCR cycles had been performed, using primers 4i5 and 4i3 (23), a 20-μl aliquot was transferred into a fresh tube containing 1 μl of 20 μM dCTP, dGTP, and dTTP and 1 μCi α-[^{33}P]dATP. 15 additional PCR cycles were then performed with both reactions, and the unlabeled PCR product was examined on an agarose gel as a control. For SSCP analysis, the labeled PCR product was digested with an appropriate enzyme to yield fragments of 150–400 bp, mixed with an equal amount of loading buffer (including 95% formamide), denatured at 95°C for 2 min, and immediately chilled on ice. The samples were then loaded onto a precooled 40-cm wedge-shaped polyacrylamide gel consisting of 5% acrylamide, 5% glycerol, and 0.8× TBE, and subjected to electrophoresis at 25 W, 10°C, for 5–6 h. The gel was fixed in 10% acetic acid/10% methanol for 30 min, dried, and exposed overnight.

cDNA analysis. Epithelial cells obtained by nasal brushing and lymphocytes collected from peripheral blood by Ficoll centrifugation

Address correspondence to Dr. Manfred Stuhrmann, Institute of Human Genetics, Medical School, 30623 Hannover, Germany.

Received for publication 29 September 1993 and in revised form 3 December 1993.

1. *Abbreviations used in this paper:* CF, cystic fibrosis; CFTR, cystic fibrosis transmembrane conductance regulator; SSCP, single strand conformation polymorphism.

J. Clin. Invest.

0021-9738/94/04/1852/08 $2.00
Volume 93, April 1994, 1852–1859

were washed with PBS and resuspended in 4 M guanidinium thiocyanate (GTC). Total RNA was extracted by the acid GTC-phenol-chloroform method as described (24). With primer CF704, 5 μg of lymphocyte RNA or the total yield of epithelial RNA were converted to cDNA using a First-Strand-cDNA-Synthesis Kit (Pharmacia, Upsala, Sweden) according to the manufacturer's protocol.

CFTR cDNA spanning exons 3–5 was amplified using the nested PCR approach depicted in Fig. 2. The sequences of PCR primers are listed in Table I. Aliquots of the final PCR were analyzed in a 2% low melting agarose gel. Bands were cut out, melted, and reamplified using the biotinylated primer B-CF297 and primer CFG178. Single strands for direct sequencing were obtained with biotin streptavidin magnetic beads (Dynal, Oslo, Norway) according to the manufacturer's protocol. Either primers CFG178 or CF740 were used for direct sequencing by the dideoxy chain termination method using the Sequenase 2.0 kit (U. S. Biochem. Corp., Cleveland, OH).

Southern blot hybridization (25). 20 μg of genomic DNA was digested using the restriction enzyme EcoRI, subjected to electrophoresis through a 0.8% agarose gel, and blotted onto N^+ Nylon membrane (Amersham Buckinghamshire, UK) using 0.4 N NaOH. The filter was hybridized with the insert DNA amplified from genomic DNA using primers 183-5′ and 183-3′ (Table I) and labeled by the random prime-labeling method (26) with α-[^{32}P]dCTP (Amersham). Overnight hybridization at 65°C was followed by two washes in 2× SSC, 0.1% SDS at room temperature, two washes at 55°C with the same solution, and a final wash with a 0.5× SSC, 0.1% SDS solution at 55°C.

Vectorette PCR (27). To analyze the 3′ end of the 183-bp cryptic exon in intron 3 of the CFTR gene, we constructed a EcoRI Vectorette library. 1 μg genomic DNA was digested for 1 h at 37°C with 20 U EcoRI (Boehringer Mannheim, Mannheim, Germany). ATP and dithiothreitol (DTT) were each added to a final concentration of ~ 2 mM. After adding 3 pmol EcoRI Vectorette unit (ICI, Abingdon, UK) and 1 U T4 DNA ligase (Boehringer Mannheim), the ligation was carried out at 16°C overnight.

An aliquot of 5 μl of the EcoRI Vectorette library was used as template for an asymmetric PCR (30 cycles) using primer 183-4. Nested PCRs using the primer combinations 183-4/vec.1 (ICI), 183-5′/vec.1, and B-183-8/vec.2 were performed (Fig. 4, Table I). After the third PCR, sequencing was performed using a combination of the biotin streptavidin magnetic beads method and the dideoxy chain termination method mentioned above.

The 3′ end sequence of the 183-bp fragment was confirmed by PCR and direct sequencing using primers B183-8 and 183-10, created from the downstream sequence of the 183-bp fragment.

Allele-specific oligonucleotide hybridization (28). 10% of the total RT-PCR (Fig. 2) was subjected to electrophoresis through a 0.8% agarose gel stained with ethidium bromide and were blotted onto N^+ Nylon membrane (Amersham) using 0.4 N NaOH. The hybridization with γ-[^{32}P]dATP-labeled oligonucleotide CFN3 (GATAGAGAG-CTGGCT) at 48°C 3 h was followed by two washes in 2× SSPE, 0.1% Triton X-100 at room temperature, and two washes at 58°C with the same solution. The x-ray film was analyzed by densitometry.

Results

The investigation of PCR products from genomic DNA of 700 German CF patients by single-strand conformation analysis and direct sequencing led to the identification of a novel nonsense mutation E92X in exon 4 of the CFTR gene (Fig. 1). This G → T transversion at the first base of exon 4 was detected in three families from Southern Germany with four CF patients. The clinical data of all four patients are summarized in Table II.

In family 1, an E92X/ΔF508 compound heterozygous female was severely affected throughout her life. Since infancy she exhibited pathognomic CF symptoms such as steatorrhea, failure to thrive, and recurrent episodes of pneumonia, but diagnosis of CF by pathological Gibson-Cooke pilocarpine iontophoresis sweat tests (90 mval Cl/liter) was delayed until the age of 4 yr 1 mo. Despite aggressive antimicrobial treatment and nutritional rehabilitation, the girl always remained below the third percentiles for height and weight. Her airways became colonized with *Pseudomonas aeruginosa* by the age of 9 yr. Amyloidosis of kidneys and liver and cor pulmonale were diagnosed when she was 16 yr old; she died of respiratory failure by the age of 17 yr 10 mo.

Table I. DNA Sequence of Oligonucleotides Used for RT and PCR

PCR primer	Sequence	
G27Art	5′ CTGGACCAGACCAATTTCCGAGG	3′
B-CF297	5′ AGAATGGGATAGAGAGCTGGCTTC	3′
CF704	5′ GTGCCAATGCAAGTCCTTCATCAA	3′
CFG178Art	5′ GGAAAGGAGACTAACAAGTTGTC	3′
CF740	5′ GCTATTCTCATCTGCATTCC	3′
183-5′	5′ CCAGCAAAAGAGATCGTAAAAGCC	3′
183-3′	5′ TCAGTAGGAGTTGAGATGATGG	3′
183-4	5′ CAGTCATGCTAAGGGCCAAA	3′
183-5	5′ GGTTTGGCCCTTAGCATGACTG	3′
183-7	5′ CCAAAGTTACACATATGGTA	3′
B-183-8	5′ GGGTGAGTGGAAAGGTACTCCTGAG	3′
183-10	5′ CTGCAGCCTGATCACACTTGTG	3′

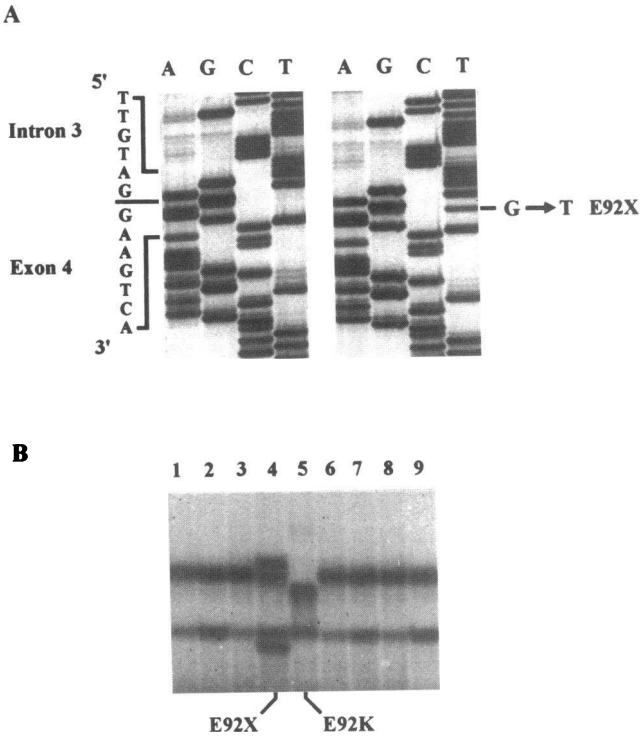


Figure 1. (*A*) DNA sequence of the CFTR exon 4/intron 3 boundary from an individual heterozygous for E92X in comparison with a wild type control. The E92X mutation is indicated by an arrow. (*B*) SSCP analysis of a HinfI fragment from exon 4 PCR products showing heterozygosity for E92X in lane *4* and homozygosity for the previously described mutation E92K (52) in lane 5, indicated by mobility shifts.

Table II. Clinical Characteristics of Patients with CF Carrying the E92X CFTR Mutation

CFTR mutation genotype	E92X/ΔF508	E92X/2789 + 5 G → A	E92X/unknown	E92X/unknown
Current age (yr/mo)	Death at 17/10	38/2	11/4	13/6
Age at diagnosis (yr)	4	20	5	7
Meconium ileus	No	No	No	No
Chloride concentration in sweat test (mmol/liter)	90	100	115	100
Pancreatic status	PI	PS	PI	PI
Onset of chronic colonization with *Pseudomonas aeruginosa* (yr)	9	35	9	9
Complications	Amyloidosis		Allergic bronchopulmonary Aspergillosis	
Current clinical data (8/1993)	(Age of 17 yr 6 mo)			
FVC (% predicted)	32	75	85	110
FEV1 (% predicted)	20	56	60	95
Chrispin Norman score of chest roentgenogram	24	20	10	6
Weight predicted for weight	75	102	98	104
Shwachman score	40	55	90	100

In contrast, CF disease has been benign in a 38-yr-old male in family 2. This patient is compound heterozygous for E92X and the splice site mutation 2789 + 5 G → A (Highsmith et al., personal communication, compiled in reference 29) on the other CF allele. The proband was asymptomatic until adolescence. Recurrent pulmonary infections at the age of 19 yr led to admission to a thorax rehabilitation clinic where CF was diagnosed by pathological sweat test values of 100 mval Cl/liter. After adequate bacteriological monitoring and intermittent antimicrobial therapy had been initiated, the patient regained excellent health. The married patient had a normal social and working life until the age of 33 yr, when his lungs became colonized with *P. aeruginosa*. Exocrine pancreatic insufficiency developed by the age of 35 yr. FVC and FEV1 have meanwhile dropped into the pathological range of 50–70% of predicted value, so that the patient is now only in part-time work and had to give up sporting activities such as alpine hiking and long-distance running.

In family 3, two siblings are compound heterozygous for E92X on the maternal allele and a different mutation on the other allele that remained unknown after genomic SSCP analysis of the whole coding region of CFTR. CF was diagnosed at the same time by pathological sweat tests (Cl^-, 100–115 mval/liter) when the patients were 5 and 7 yr old, respectively. The 11- and 13-yr-old girls are pancreas insufficient, have normal lung function, and have always exhibited growth beyond the 90th percentiles for height and weight. Both patients acquired *P. aeruginosa* in their airways by the age of 9 yr. With the exception of episodes of allergic bronchopulmonary aspergillosis in the younger girl, CF disease has so far been benign in the siblings (actual Shwachman scores, > 80, stage I).

In all three families, the E92X chromosome is characterized by an uncommon combination of the extragenic upstream marker haplotype 1-2-2-1 (XV2c-KM.19-MP6D9-J44) and the intragenic CFTR marker haplotype 1-2-2-1-1 (GATT-TUB9-M470V-T854-TUB18), which could have evolved by recombination from the two most common dimorphic marker haplotypes in German families (19). This infrequent haplotype was not seen in other German CF chromosomes in our panel, suggesting that it could be diagnostic for the E92X mutation probably due to a founder effect. This hypothesis is further strengthened by the common geographic origin of all three families from the same small area of Southern Germany. As expected from the haplotype data, an additional screening for E92X with the restriction enzyme EcoNI did not uncover further cases among 800 German CF chromosomes tested this way.

Since the E92X mutation affects the first nucleotide of exon 4, we surmised a possible effect on the splicing of CFTR pre-mRNA. We thus analyzed CFTR transcripts from epithelial cells of the two E92X heterozygous siblings and lymphocytes of these two sisters and their mother who transmitted this mutation. Due to the low amount of CFTR transcript in nonspecific tissue such as lymphocytes and because of the low yield of epithelial cells from nasal brushing, we used nested RT-PCR with primers spanning exons 2–6 of the CFTR cDNA.

From lymphocytes of all three probands, we obtained a characteristic product pattern with two bands of 391 and 175 bp (Fig. 2). Direct sequencing of the excised RT-PCR product from all three probands showed the wild type sequence without the E92X mutation in the 391-bp product, thus demonstrating skipping of exon 4 in the 175-bp product (Fig. 3). This exon skipping was neither observed in 10 further lymphocyte samples from healthy control persons nor in the paternal lymphocyte sample. A different pattern was obtained from nasal epithelial cells of the patients and their mother, which gave an additional product of 574 bp (Fig. 2). We performed ASO hybridization using oligonucleotide CFN3 as a probe to quantitate the relative amount of the different PCR products by densitometry (nasal epithelial cells: 574:391:175 bp, 0.8:1:0.2; lymphocytes: 391:175:, 1:0.94). Direct sequencing of the 574-bp product revealed that a 183-bp sequence was inserted into the CFTR cDNA within the junction of exons 3 and 4 (Fig. 3 *B*). Neither this additional sequence nor exon 4 skipping was observed in the paternal nasal epithelial sample and in further nasal epithelial samples from six non-E92X CF patients and eight unrelated non-CF control persons. Furthermore, the

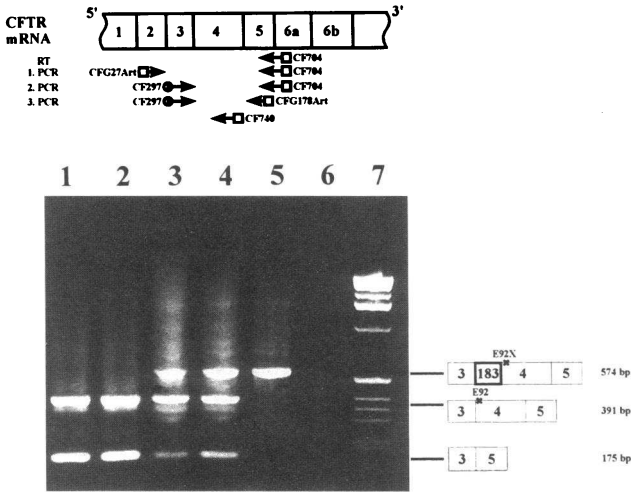


Figure 2. Schematic representation of the CFTR cDNA from exon 1 to exon 6b. RT-PCR analysis of CFTR transcript spanning exons 2–6a was performed by three rounds of PCR with the indicated pairs of nested oligonucleotide primers. The products were separated by 2% agarose gel electrophoresis. RT-PCR products from lymphocytes (lanes *1* and *2*) and nasal epithelial cells (lanes *3* and *4*) of the CF patients of family 3 consisted of fragments of the expected size (391 bp, lanes *1–4*), fragments lacking exon 4 (175 bp, lanes *1–4*), and fragments including a 183-bp sequence (574 bp, lanes *3* and *4*). Lane *5*, the larger fragment after isolation and reamplification: lane *6*, RT-PCR without mRNA; lane *7*, DNA size marker (1-kb ladder; GIBCO BRL, Gaithersburg, MD).

larger RT-PCR product was neither found in pancreas or other previously analyzed (21) control tissues from non-CF individuals, nor in the colon cell lines HT-29 or T84.

The aberrantly spliced mRNA was found to be transcribed only from the E92X allele (Fig. 3 *D*) and contains two in-frame nonsense codons (Fig. 4 *C*). We confirmed the human origin of the 183-bp insert by PCR from genomic DNA using primers located within the inserted sequence (183-5′ and 183-3′). Assuming the 183-bp insert to be the result of a splicing event within intron 3 of the CFTR gene, PCR with primers specific to exon 3 and the insert was performed from genomic DNA, resulting in amplification of a 2.7-kb fragment. Direct sequencing of this PCR product showed the 183-bp sequence to be located ~ 2.2 kb downstream of exon 3 and to be preceded by an acceptor splice site with a short pyrimidine stretch (Fig. 4 *A*). With additional sequencing primers, we obtained further sequence information from ~ 300 bp of the 3′ splice site preceding the inserted sequence (Fig. 4 *C*). Since most branch points are located within 18–40 bp from the 3′ splice junction (30), two possible branch points are favored (Fig. 4 *C*). Interestingly, an apparently more perfect branch point sequence is located further upstream of the new 183-bp exon followed by another apparently silent 3′ splice site with a longer polypyrimidine stretch. In addition, this part of intron 3 seems highly repetitive.

For the determination of downstream sequences, we used Vectorette PCR as recently described (27). Southern blot analysis of genomic DNA using the insert sequence as a probe guided the choice of the appropriate Vectorette unit. Since the 183-bp insert hybridized to an ~ 4.5-kb EcoRI fragment, we constructed a genomic EcoRI Vectorette library and conducted nested PCR with primers specific to the 183-bp exon and the Vectorette unit (Fig. 4 *B*). Direct sequencing of the final PCR product revealed a perfect donor splice site flanking the 183-bp exon (Fig. 4 *C*). We then established intron primers for the amplification of the whole new exon and its flanking intron sequences (Table I). SSCP analysis and direct sequencing of genomic DNA from the patients and 10 further individuals harboring different CFTR haplotypes did not show any polymorphisms or sequence variants in the analyzed region.

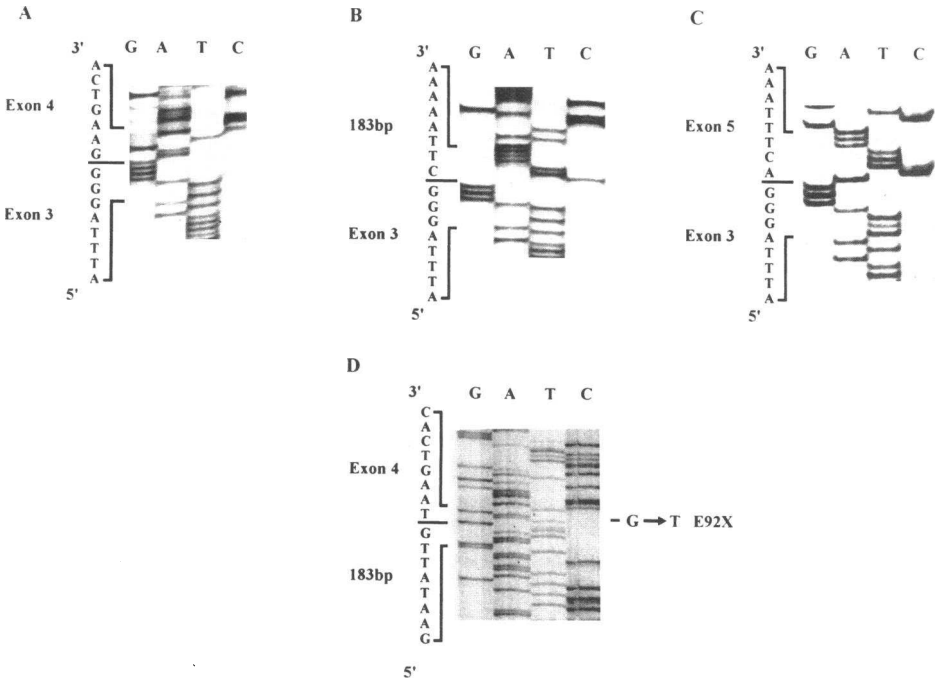
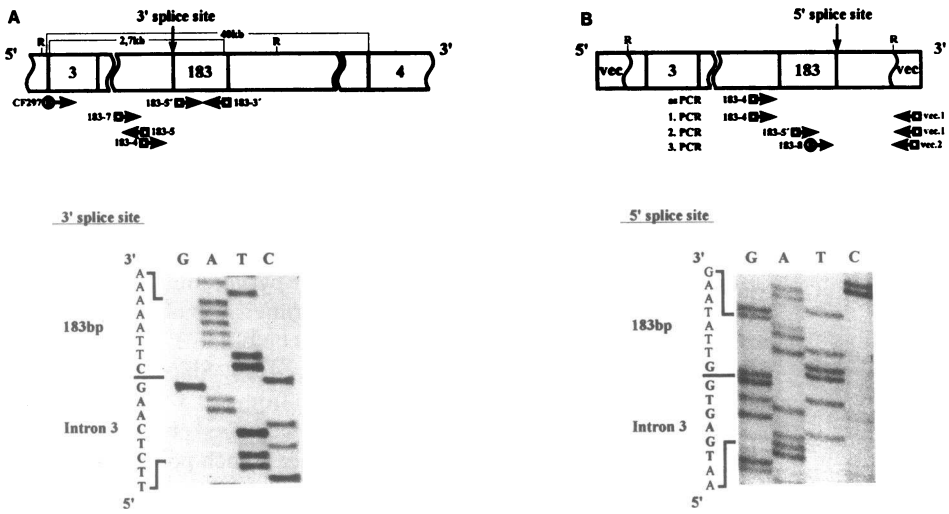


Figure 3. Direct sequencing of the RT-PCR products of CFTR cDNA reamplified with primers B-CF297 and CFG178Art. (*A*) Normal-sized product (391 bp), without the E92X mutation within exon 4 at position +1; (*B*) 5′ boundary of the 183-bp fragment within the 574-bp products; (*C*) the 175-bp product, missing exon 4; and (*D*) 3′ boundary of the 183-bp fragment within the 574-bp product, showing the E92X mutation.



C

```
5'
AACAGAAGTTGAGCGTGAGGCAAATTAAGTAACTTTCCCAAAGTTACACATATGGTAAGT

TTGTGAGATATCCCAGTCTCTTTAGCTCCAAAGCCTTTGACCCTTTCACCATACCAGATT

ATGATTCGTATTAATATATAATTATAATTATAATGATTGTATTTAGGTACTCAACAGAAT

GGTGACTCTAGTAACCAGCCTTGGTTCTGCTGAGCTTCTCTGCGTCTTCTCAGGAGACAC

AGGCTACAGAGCTTGAAGGCTGAGGATTCTTCCAGGGTCACTTCAGGGGCAAATCTGAAA

CTTTCTTCAGGACAGGAATCAACGAGATCTTCTCACTTACTTATACCTGGGGGAGGAACT

GTATGAATCCACCCAAGAACCAGTCATGCTAAGGGCCAAACCAGTCATGCTAAGGGCCAA

ACCTATAGACAAAAAAGGGATAGGAGAATGGAGTATGTATGGAGAAAGACTAAATTGTTC
                L  K  N  I  P  A  K  E  I  V  K  A  L  H  G
TTAAACTTCTCAAG[CTTAAAAATATCCCAGCAAAAGAGATCGTAAAAGCCCTTCATGGC
 V  L  I  I  H  A  W  G
GTATTAATTATCCATGCATGGGGGTGAGTGGAAAGGTACTCCTGAGCCCGAGGCTACAGC

TTTGGAACTAGCAGCACCTTTGAAGGGGAAAGCGTGTTTCCATCATCTCAACTCCTACTG

ATAACCAATGGAATATTG]GTGAGTAAAGGATCCTGGGGGAAGAAGCAGCTGAAATGTGT

AGGTGAGAAGGCAGAGAGAAGAATATTTATATTGGGAATGGCACAAGTGTGATCAGGCTG

CAGGTTTTTCATCCTTGTCATAGAGAAAAAACCACGCTGGACA
                                                3'
```

Figure 4. Direct genomic sequencing of the 183-bp fragment and its flanking intron 3 sequences. Schematic map of the location of the 183-bp fragment and the oligonucleotide primers, which were used for the splice site identification. (*A*) Identification of the 3′ splice site and a potential branch point. The intron 3 sequence between exon 3 and the 183-bp fragment was amplified by PCR with primer 183-3′ and biotinylated primer B-CF297. Direct sequencing of this PCR product using a combination of the biotin streptavidin magnetic beads method and the dideoxy chain termination method with primer 183-3′ revealed the 3′ splice site sequence. The potential branch point was searched by PCR using an intron 3 primer (183-4, 183-5, or 183-7) in combination with primers B-CF297 and 183-3′, respectively. (*B*) Characterization of the 5′ splice site of the 183-bp fragment. The 183-bp fragment hybridized to an ~ 4.5-kb EcoRI genomic fragment. Its boundaries are indicated (*R*). To identify the intron 3 sequence downstream of the 183-bp cryptic exon, a genomic EcoRI Vectorette library was constructed and used as template for nested PCRs with primer pairs specific for the 183-bp fragment (183-4, 183-5′, B183-8) and the Vectorette sequence (vec.1, vec.2). Direct sequencing of the final PCR product uncovered a perfect 5′ splice site at the intron 3/183-bp boundary. (*C*) Sequence of the 183-bp fragment and its surroundings in intron 3 of the CFTR gene. The 183-bp fragment (enclosed by *brackets*), containing two stop codons (*bold letters, underlined*), is flanked by perfect 5′ and 3′ splice sites with a short oligopyrimidine stretch (*bold letters*). Two candidate sequences for the branch point are underlined. A silent 3′ splice site with a large polypyrimidine stretch and a perfect branch point is located further upstream of the 183-bp exon (*underlined with dashes*). The deduced amino acid sequence for the cryptic exon (until the stop codon) is shown above the nucleotide sequence in a single-letter code. These data were submitted to the LANL Database under accession number L25269.

Finally, we found no sequence homology of this cryptic exon and the deduced peptide to other genes or proteins, respectively.

Discussion

The clinical course of disease in CF patients is highly variable. It has, however, been shown that the onset of many clinical symptoms, including pancreatic insufficiency and airways' colonization with opportunistic pathogens, is partly determined by the CFTR mutation genotype (31–33), which appears to be the main factor in the genetic predisposition for CF. CFTR mutations have been classified according to their effect on pancreatic function as pancreas insufficient ("severe") or pancreas sufficient ("mild") alleles, and by this definition the maintenance of pancreatic function in CF requires the presence of at least one pancreas-sufficient allele (32, 33). Following that classification, the E92X mutation is a pancreas-insufficient allele, because the E92X/ΔF508 patient was pancreas insufficient since infancy. The second E92X patient with completely known CFTR genotype and a less severe clinical phenotype carries a pancreas sufficient mutation, 2789 + 5 G → A, on his other chromosome.

CF with E92X is the first case of a human inherited disease, as far as we know, in which the consequences of a change of the first nucleotide of an exon were studied in patients' mRNA transcripts. This G → T transversion affects the 3′ splice site of exon 4 and furthermore creates an in-frame stop codon. RT-PCR analyses indicated the ability of this mutation to induce exon 4 skipping of the transcripts (lymphocytes and nasal epithelial cells) as well as activation of a cryptic splice site in intron 3, resulting in the inclusion of an additional 183-bp fragment.

The 183-bp sequence was present in nasal epithelial cell–derived mRNA of the two E92X heterozygous probands. As shown by direct sequencing (Fig. 3 *D*), it was only transcribed from the E92X allele. There was no evidence for use of the cryptic splice sites in CFTR transcripts of lymphocytes from patients and several controls. Furthermore, the cryptic exon has never been detected in kidney, liver, jejunum, placenta, lung, and nasal polyps from non-CF controls (21).

The differences in the splicing pattern of E92X transcripts, identical in all analyzed specimens, point to a tissue-specific splicing event. Since differential expression of various *cis*-acting elements and *trans*-acting factors have been reported to be involved in determination of tissue or developmental stage specificity (34–36), it is not unexpected that a mutation of the 3′ splice site sequence may lead to different splicing patterns in ectopic and expressing tissues. However, no further cases of tissue-specific splicing of CFTR transcripts are known to date.

Normally spliced CFTR transcripts carrying the E92X mutation were not observed in both tissues. Several reports of severely reduced mRNA levels as a consequence of nonsense mutations in the CFTR gene have been published (7–9, 11), but in case of the E92X mutation, densitometric quantification of ASO-hybridized PCR products suggests that the main effect of this mutation is aberrant splicing. However, we cannot completely exclude normally spliced but reduced proportions of transcripts from E92X-carrying alleles, because the yield of PCR products could have been biased by the cDNA size in nested PCR.

The observed in-frame-skipping of exon 4 and the severity of the E92X allele would be consistent with a proposed important role of exon 4 for CFTR function. Skipping of exon 4 from CFTR transcripts has been reported in humans (21) and mice (37), but exon 4^- transcripts represented only a very small proportion of total CFTR mRNA from human tissues (21), and did not appear to be associated with clinical symptoms. Exon skipping does not always cause disease; nonsense mutation–induced exon skipping has even been shown to function as a mechanism for phenotypic rescue of the common inherited defect of the AMPD1 (AMP deaminase) gene (38). Such a phenotypic rescue in case of the E92X mutation can probably not be realized by omission of exon 4, given the severe clinical course in the ΔF508/E92X compound heterozygous patient, the high number of other known disease-causing CF mutations affecting this exon (29), and the postulated role of exon 4 in coding for a large part of the CFTR transmembrane pore (3). Even if residual functions of exon 4^- CFTR have not been excluded by in vitro experiments so far, parallels can be drawn from the frequent splice site mutation 621 + 1 G → T, which also leads to skipping of exon 4 (13, 39). A comprehensive genotype–phenotype study of patients carrying the splice mutation 621 + 1 G → T strongly argues that the extensive exon 4 skipping is associated with an early onset of severe clinical symptoms of CF (40). In the E92X patients, the fraction of exon 4 skipping has been relatively low in epithelial cells compared with the proportion of the in-frame cryptic exon inclusion. The inserted cryptic exon, however, by itself introduces two new nonsense codons into the E92X transcript, and its translation would lead to a severely truncated protein containing only 7% of wild type CFTR.

Cryptic exons have already been identified in other regions of the CFTR gene (41–43, and Highsmith et al., personal communication, compiled in reference 29). Cryptic 119- and 260-bp exons in intron 11 and 23, respectively, were observed in epithelial and nonepithelial cells from CF and non-CF probands (41, 42). Both alternatively spliced transcripts occurred at a low level without any obvious phenotypic consequences. The physiological significance of two other recently identified exons upstream of CFTR exon 1 (exon 1a and exon −1a; reference 43) is unclear. The aberrant activation of a cryptic exon in intron 19, however, causes CF disease (Highsmith et al., personal communication, compiled in reference 29). The underlying single base pair substitution (3849 + 10-kb C → T) is usually associated with relatively benign CF in patients who are compound heterozygous for this mutation and ΔF508 (44), thus determining the corresponding allele as being pancreas sufficient, in contrast to the E92X allele. A possible explanation for this discrepancy is that the 3849 + 10-kb C → T mutation leaves some wild type function with normally spliced transcript, whereas we have found E92X to be strictly associated with aberrant splicing. Of general interest is the sequence position of the G → T transversion described here. It affects the first nucleotide of exon 4 of the CFTR gene. Guanosine is the most common base at this position of the 3′ splice site, while, within the CFTR gene, only exon 17a begins with thymidine (23). Although several disease-causing mutations in 3′ splice sites are known to result in exon skipping and/or in activation of cryptic splice sites, for example in β-thalassaemia or LPL deficiency (reviewed in reference 45), a G → T transversion at the first nucleotide of an exon has so far only been studied by site-directed mutagenesis in an artificial hamster dihydrofolate reductase minigene (46). In this CHO model system, exon skipping was observed in 52% of the transcripts. In our study, we demonstrate that a G → T transversion at this position within the splice acceptor site can result in both exon skipping and the recruitment of a cryptic exon, in vivo. The detection of these splice effects cannot be explained completely by a simple comparison of the sequences of the altered splice acceptor site and the cryptic splice site within intron 3 with the consensus sequence, since another cryptic splice site located immediately upstream of the 183-bp sequence has even more favorable statistical scores (45, 47). Furthermore, it is noteworthy that the E92X allele was found exclusively in conjunction with the cryptic 183-bp exon. Exon 3b was not seen in normal controls, and E92X did not appear within a normally spliced product. A possible interpretation of this observation is that both the donor splice site of cryptic exon 3b and the weakened acceptor splice site with E92X are only used if they can be recognized together, and this recognition appears to be tissue specific. Alternatively, epithelial cells may generally use the cryptic exon in a first splicing step, but its subsequent removal may be disturbed by the presence of E92X within the acceptor splice site. The exon sequence adjacent to the invariant consensus sequence has been shown to be a determinant for splice site selection in natural pre-mRNAs as well as in artificial constructs (48, 49), and consequently exon mutations as E92X can change the relative strength of competing splice sites, resulting in aberrant splicing. Furthermore, a single base mutation can have long-range effects on the concerted splicing of other exons, leading to multiple exon skipping (50). Besides several *cis*-acting elements, various antagonistic splicing factors such as SF2 and ASF, and heterogeneous nuclear ribonucleoproteins such as hnRNP A1 appear to be involved in the tissue

specificity of splicing (51). Future characterization of *cis*- and *trans*-acting factors and of their spatial and temporal interaction will improve our understanding of the role of splice site selection in differentiation and disease.

Acknowledgments

We are grateful to the patients and their families who provided samples for this study. We thank Brigitte Wulf for technical assistance, Dr. Michael Krawczak for helpful discussions, Dr. David N. Cooper for reading the manuscript, and Professor Horst von der Hardt and Professor Günther Maaß for continuous support and encouragement.

Grants from the Deutsche Forschungsgemeinschaft are gratefully acknowledged. K.W. is a fellow of the Deutsche Gesellschaft zur Bekämpfung der Mukoviszidose (Cystic Fibrosis Fund of Germany).

References

1. Boat, T. F., M. J. Welsh, and A. L. Beaudet. 1989. Cystic fibrosis. *In* The Metabolic Basis of Inherited Disease. 6th ed. C. L. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, editors. McGraw-Hill Inc., New York. 2649–2680.

2. Frizell, R. A., G. Rechkemmer, and R. L. Shoemaker. 1986. Altered regulation of airway epithelial cell chloride channels in cystic fibrosis. *Science (Wash. DC)*. 233:558–560.

3. Riordan, J. R., J. M. Rommens, B.-S. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic, J. Chou, et al. 1989. Identification of the cystic fibrosis gene: cloning and characterization of complementary cDNA. *Science (Wash. DC)*. 245:1066–1073.

4. Rommens, J. M., M. L. Iannuzzi, B.-S. Kerem, M. L. Drumm, G. Melmer, M. Dean, R. Rozmahel, J. L. Cole, D. Kennedy, N. Hidaka, et al. 1989. Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science (Wash. DC)*. 245:1059–1065.

5. Kartner, N., J. W. Hanrahan, T. J. Jensen, A. L. Naismith, S. Sun, C. A. Ackerley, E. F. Reyes, L. C. Tsui, J. M. Rommens, C. E. Bear, and J. R. Riordan. 1991. Expression of the cystic fibrosis gene in non-epithelial invertebrate cells produces a regulated anion conductance. *Cell*. 64:681–691.

6. Kerem, B.-S., J. M. Rommens, J. A. Buchanan, D. Markiewicz, T. K. Cox, A. Chakravarti, M. Buchwald, and L.-C. Tsui. 1989. Identification of the cystic fibrosis gene: genetic analysis. *Science (Wash. DC)*. 245:1073–1080.

7. Hamosh, A., B. C. Trapnell, P. L. Zeitlin, C. Montrose-Rafizadeh, B. J. Rosenstein, R. G. Crystal, and G. R. Cutting. 1991. Severe deficiency of cystic fibrosis transmembrane conductance regulator messenger RNA carrying nonsense mutations R553X and W1316X in respiratory epithelial cells of patients with cystic fibrosis. *J. Clin. Invest.* 88:1880–1885.

8. Jones, C., I. McIntosh, M. Keston, A. Ferguson, and D. J. H. Brock. 1992. Three novel mutations in the cystic fibrosis gene detected by chemical cleavage: analysis of variant splicing and a nonsense mutation. *Hum. Mol. Genet.* 1:11–17.

9. Hamosh, A., B. J. Rosenstein, and G. R. Cutting. 1992. CFTR nonsense mutations G542X and W1282X associated with severe reduction of CFTR mRNA in nasal epithelial cells. *Hum. Mol. Genet.* 1:542–544.

10. Smit, L., S. Z. Nasr, M. C. Iannuzzi, and F. S. Collins. 1993. An African-American cystic fibrosis patient homozygous for a novel frameshift mutation associated with reduced CFTR mRNA levels. *Hum. Mut.* 2:148–151.

11. Will, K., J. Reiss, M. Dean, M. Schlösser, R. Slomski, J. Schmidtke, and M. Stuhrmann. 1993. CFTR transcripts are undetectable in lymphocytes and respiratory epithelial cells of a CF patient homozygous for the nonsense mutation R553X. *J. Med. Genet.* 30:833–837.

12. Fonknechten, N., J.-C. Chomel, A. Kitzis, A. Kahn, and J.-C. Kaplan. 1992. Skipping of exon 5 as a consequence of the 711 + 1 G → T mutation in the CFTR gene. *Hum. Mol. Genet.* 1:281–282.

13. Hull, J., S. Shackleton, and A. Harris. 1993. Abnormal mRNA splicing resulting from three different mutations in the CFTR gene. *Hum. Mol. Genet.* 2:689–692.

14. Dörk, T., K. Will, A. Demmer, and B. Tümmler. 1993. A donor splice mutation (405 + 1 G → A) in cystic fibrosis associated with exon skipping in epithelial CFTR mRNA. *Hum. Mol. Genet.* 2:1965–1966.

15. Cheng, S. H., R. J. Gregory, J. Marshall, S. Paul, D. W. Souza, G. A. White, C. R. O'Riordan, and A. E. Smith. 1990. Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell*. 63:827–834.

16. Gregory, R. J., D. P. Rich, S. H. Cheng, D. W. Souza, S. Paul, P. Manavalan, M. P. Anderson, M. J. Welsh, and A. E. Smith. 1991. Maturation and function of cystic fibrosis transmembrane conductance regulator variants bearing mutations in putative nucleotide-binding domains 1 and 2. *Mol. Cell. Biol.* 11:3886–3893.

17. Dalemans, W., P. Barby, G. Champigny, S. Jallat, K. Dott, D. Dreyer, R. G. Crystal, A. Pavirani, J.-P. Lecocq, and M. Lazdunski. 1991. Altered chloride ion channel kinetics associated with the ΔF508 cystic fibrosis mutation. *Nature (Lond.)*. 354:526–528.

18. Sheppard, D. N., D. P. Rich, L. S. Ostedgaard, R. J. Gregory, A. E. Smith, and M. J. Welsh. 1993. Mutations in CFTR associated with mild disease form Cl^- channels with altered pore properties. *Nature (Lond.)*. 362:160–164.

19. Dörk, T., T. Neumann, U. Wulbrand, B. Wulf, N. Kälin, G. Maaß, M. Krawczak, H. Guillermit, C. Ferec, G. Horn, et al. 1992. Intra- and extragenic marker haplotypes of CFTR mutations in cystic fibrosis families. *Hum. Genet.* 88:417–425.

20. Ferrie, R. M., M. J. Schwarz, N. H. Robertson, S. Vaudin, M. Super, G. Malone, and S. Little. 1992. Development, multiplexing, and application of ARMS tests for common mutations in the CFTR gene. *Am. J. Hum. Genet.* 51:251–262.

21. Bremer, S., T. Hoof, M. Wilke, R. Busche, B. Scholte, J. R. Riordan, G. Maass, and B. Tümmler. 1992. Quantitative expression patterns of multidrug-resistance P-glycoprotein (MDR1) and differentially spliced cystic fibrosis transmembrane conductance regulator mRNA transcripts in human epithelia. *Eur. J. Biochem.* 206:137–149.

22. Orita, M., H. Iwahana, K. Kanazawa, and T. Sekiya. 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-stranded conformation polymorphisms. *Proc. Natl. Acad. Sci. USA*. 86:2766–2770.

23. Zielenski, J., R. Rosmahel, D. Bozon, B.-S. Kerem, Z. Grzelczak, J. R. Riordan, J. M. Rommens, and L.-C. Tsui. 1991. Genomic DNA sequence of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. *Genomics*. 10:214–228.

24. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156–159.

25. Southern, E. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503–517.

26. Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction fragments to high specific activity. *Anal. Biochem.* 132:6–13.

27. Arnold, C., and I. J. Hodgson. 1991. Vectorette PCR: a novel approach to genomic walking. *PCR Methods. Applic.* 1:39–42.

28. Saiki, R. L., T. L. Bugawan, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1986. Analysis of enzymatically amplified β-globin and HLA-DQα DNA with allele-specific oligonucleotide probes. *Nature (Lond.)*. 324:163–166.

29. Tsui, L.-C. 1992. Mutations and sequence variations detected in the cystic fibrosis transmembrane conductance regulator (CFTR) gene: a report from the cystic fibrosis genetic analysis consortium. *Hum. Mut.* 1:197–203.

30. Reed, R. 1989. The organization of 3′ splice-site sequences in mammalian introns. *Genes & Dev.* 3:2113–2123.

31. Kerem, E., M. Corey, B.-S. Kerem, J. M. Rommens, D. Markiewicz, H. Levison, L.-C. Tsui, and P. Durie. 1990. The relation between genotype and phenotype in cystic fibrosis-analysis of the most common mutation (ΔF508). *N. Engl. J. Med.* 323:1517–1522.

32. Kristidis, P., D. Bozon, M. Corey, D. Markiewicz, J. Rommens, L.-C. Tsui, and P. Durie. 1992. Genetic determination of exocrine pancreatic function in cystic fibrosis. *Am. J. Hum. Genet.* 50:1178–1184.

33. Kubesch P., T. Dörk, U. Wulbrand, N. Kälin, T. Neumann, B. Wulf, H. Geerlings, H. Weißbrodt, H. von der Hardt, and B. Tümmler. 1993. Genetic determinants of cystic fibrosis airways' colonization with Pseudomonas aeruginosa. *Lancet*. 341:189–193.

34. Libri D., L. Balvay, and M. Y. Fiszman. 1992. In vivo splicing of the β tropomyosin pre-mRNA: A role for branch point and donor site competition. *Mol. Cell. Biol.* 12:3204–3215.

35. Bingham, P. M., T.-B. Chou, I. Mims, and Z. Zachar. 1988. On/off regulation of gene expression at the level of splicing. *Trends Genet.* 4:134–138.

36. Valcarcel J., R. Singh, P. D. Zamore, and M. R. Green. 1993. The protein sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of *transformer* pre-mRNA. *Nature (Lond.)*. 362:171–175.

37. Delaney, S. J., D. P. Rich, S. A. Thomson, M. R. Hargrave, P. K. Lovelock, M. J. Welsh, and B. J. Wainwright. 1993. Cystic fibrosis transmembrane conductance regulator splice variants are not conserved and fail to produce chloride channels. *Nature Genetics*. 4:426–431.

38. Morisaki, H., T. Morisaki, L. K. Newby, and E. W. Holmes. 1993. Alternative splicing: A mechanism for phenotypic rescue of a common inherited defect. *J. Clin. Invest.* 91:2275–2280.

39. Zielenski, J., D. Bozon, D. Markiewicz, G. Aubin, F. Simard, J. M. Rommens and L.-C. Tsui. 1993. Analysis of CFTR transcripts in nasal epithelial cells and lymphoblasts of a cystic fibrosis patient with 621 + 1 G → T and 711 + 1 G → T mutations. *Hum. Mol. Genet.* 2:683–687.

40. The Cystic Fibrosis Genotype-Phenotype Consortium. 1993. Correlation between genotype and phenotype in patients with cystic fibrosis. *N. Engl. J. Med.* 329:1308–1313.

41. Will, K., M. Stuhrmann, M. Dean, and J. Schmidtke. 1992. Alternative splicing in the first nucleotide binding fold of CFTR. *Hum. Mol. Genet.* 2:231–235.

42. Yoshimura, K., C.-S. Chu, and R. G. Crystal. 1993. Alternative splicing of intron 23 of the human cystic fibrosis transmembrane conductance regulator gene resulting in a novel exon and transcript coding for a shortened intracytoplasmic C terminus. *J. Biol. Chem.* 268:686–690.

43. Koh, J., T. J. Sferra, and F. S. Collins. 1993. Characterization of the cystic fibrosis transmembrane conductance regulator promoter region. *J. Biol. Chem.* 268:15912–15921.

44. Augarten, A., B.-S. Kerem, Y. Yahav, S. Noiman, Y. Rivlin, A. Tal, H. Blau, L. Ben-Tur, A. Szeinberg, E. Kerem, and E. Gazit. 1993. Mild cystic fibrosis and normal or borderline sweat test in patients with the 3849 + 10Kb C → T mutation. *Lancet.* 342:25–26.

45. Krawczak, M., J. Reiss, and D. N. Cooper. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum. Genet.* 90:41–54.

46. Chen, I.-T., and L. A. Chasin. 1993. Direct selection for mutations affecting specific splice sites in a hamster dihydrofolate reductase minigene. *Mol. Cell. Biol.* 13:289–300.

47. Shapiro, M. B., and P. Senapathy. 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* 15:7155–7174.

48. Robberson, B. L., G. J. Cote, and S. M. Berget. 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* 10:84–94.

49. Talerico, M., and S. M. Berget. 1990. Effect of 5′ splice site mutations on splicing of the preceding intron. *Mol. Cell. Biol.* 10:6299–6305.

50. Carothers, A. M., G. Urlaub, D. Grunberger, and L. A. Chasin. 1993. Splicing mutants and their second-site suppressors at the dihydrofolate reductase locus in chinese hamster ovary cells. *Mol. Cell. Biol.* 13:5085–5098.

51. Mayeda, A., D. M. Helfman, and A. R. Krainer. 1993. Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. *Mol. Cell. Biol.* 13:2993–3001.

52. Nunes, V., M. Chillon, T. Dörk, B. Tümmler, T. Casals, and X. Estivill. 1993. A new missense mutation (E92K) in the first transmembrane domain of the CFTR gene causes a benign cystic fibrosis phenotype. *Hum. Mol. Genet.* 2:79–80.