Am. J. Hum. Genet. 47:611–615, 1990

Benign Missense Variations in the Cystic Fibrosis Gene

Keiko Kobayashi,* Michael R. Knowles,† Richard C. Boucher,† William E. O'Brien,* and Arthur L. Beaudet*

*Howard Hughes Medical Institute and Institute for Molecular Genetics, Baylor College of Medicine, Houston; and †Department of Medicine, University of North Carolina, Chapel Hill

Summary

The common mutation causing cystic fibrosis is a deletion of phenylalanine 508 (ΔF508), which occurs in a putative nucleotide-binding fold of the gene product. We report two additional mutations, substitution of cysteine for phenylalanine 508 (F508C) and substitution of valine for isoleucine 506 (I506V). Three compound heterozygous persons, two ΔF508/F508C and one ΔF508/I506V, had normal clinical and epithelial physiological studies indicating that the F508C and I506V mutations are benign. This opportunity to study the in vivo function of these mutations suggests that amino acid substitutions are more benign than changes in the length of this portion of the putative nucleotide-binding fold. These mutations must be taken into account when performing molecular diagnosis and carrier detection for cystic fibrosis.

Introduction

Cystic fibrosis (CF) is a well-known autosomal recessive disorder affecting between one in 2,000 and one in 3,000 individuals in the Caucasian population (Boat et al. 1989). The disease affects epithelial cells in the airways, pancreas, intestine, and sweat glands, leading to chronic lung disease, pancreatic insufficiency, and increased concentration of sodium and chloride in sweat. The clinical, electrophysiological, and genetic aspects of the disease were reviewed recently (Boat et al. 1989). An intensive effort using genetic linkage information ultimately led to the cloning of the CF gene prior to the identification of the gene product or its function (Kerem et al. 1989; Riordan et al. 1989; Rommens et al. 1989). The gene encodes what is believed to be a transmembrane protein, which has been named the cystic fibrosis transmembrane conductance regulator (CFTR). The CFTR contains two nucleotide-binding regions which show homology to numerous transport proteins with the greatest homology to the P-glycoproteins that are encoded by the multiple drug-resistance loci (Kerem et al. 1989; Riordan et al. 1989).

A three-base deletion resulting in the loss of phenylalanine residue 508 (designated ΔF508) is the mutation occurring on the majority of CF chromosomes (Kerem et al. 1989; Riordan et al. 1989), being found on 70%–75% of North American CF chromosomes (Kerem et al. 1989; Riordan et al. 1989; Lemna et al. 1990). During analysis of over 200 North American CF families for the ΔF508 mutation, we identified three parents who initially appeared to be homozygous for the ΔF508 mutation (Lemna et al. 1990). We have now determined that these three "index parents" are asymptomatic compound heterozygotes carrying the ΔF508 mutation on one chromosome and a missense mutation in the same region on the other chromosome.

Material and Methods

DNA Analysis

Genomic DNA was prepared and RFLPs were analyzed as described elsewhere (Lemna et al. 1990). For XV-2c and KM-19, the 1 allele indicates the absence of the polymorphic restriction-enzyme site and the 2 allele indicates the presence of the restriction-enzyme site. Haplotypes ([XV-2c]–[KM-19]) are as follows: A = 1–1, B = 1–2, C = 2–1, and D = 2–2. The poly-

Received March 27, 1990; revision received June 13, 1990.

Address for correspondence and reprints: Arthur L. Beaudet, M.D., Howard Hughes Medical Institute, Baylor College of Medicine, Houston, TX 77030.

0002-9297/90/4704-0004$02.00

merase chain reactions (PCR) were performed in a final volume of 50 μl containing 6.7 mM $MgCl_2$, 67 mM Tris HCl pH 8.8, 1.5 mM of each dNTP, 0.5 μM of each PCR primer, 500 ng of genomic DNA, and one unit of *Taq* polymerase (Perkin-Elmer/Cetus, Norwalk, CT). The C16B and C16D primers (Kerem et al. 1989) were modified to add restriction-enzyme sites as follows: 5'-GCTAGGTACCTCCTGGATTATGCCTGGCAC-3' and 5'-GCTACTGCAGGTTGGCATGCTTTGATGACG-3' (the underlined sequences are from the CF gene). Reactions were heated to 94°C for 5 min and subjected to 30 cycles of denaturation at 94°C for 1 min, annealing at 60°C for 45 s, and DNA polymerization at 72°C for 2 min. After amplification, a portion of the reaction mixture was analyzed by 3% agarose and/or 10% acrylamide-gel electrophoresis to assess the size of the amplified sequence and/or heteroduplex formation.

For cloning and sequencing, the amplified product was treated with phenol/chloroform, collected by ethanol precipitation, and digested with *Kpn*I and *Pst*I. The specific fragment to be cloned was purified from low-melting-point agarose following electrophoresis. DNA fragments were directionally cloned into M13mp19 which had been digested with *Kpn*I and *Pst*I. Single-stranded DNA cloned into M13 was sequenced by the dideoxy chain termination method (Biggin et al. 1983) using Sequenase and universal primer (United States Biochemical, Cleveland). A total of 49 M13 clones were sequenced from the three index parents and from the three grandparents carrying the unusual allele. Hybridization with allele-specific oligonucleotides was performed as described elsewhere (Lemna et al. 1990), except for the following variations. The allele-specific oligonucleotides were as follows: normal = 5'-CACCAAAGATGATATTTTC-3', ΔF508 = 5'-AACACCAATGATATTTTCTT-3', F508C = 5'-CACCACAGATGATATTTTC-3', and I506V = 5'-CAAAGATGACATTTTCTTT-3'. Filters were washed in 1 liter of 40 mM sodium phosphate buffer pH 7.2 containing 1% SDS for 5 min at room temperature followed by a second wash in 500 ml for 15 min at 42°C using prewarmed solution. Filters were exposed using an intensifying screen for 2–3 h.

Physiological Studies

Sweat chloride was measured by pilocarpine iontophoresis (Gibson and Cooke 1959). The cyclic AMP–regulated sweat acinar Cl^- secretion was assessed by measuring the sweat rate in response to intradermal injection of a mixture of isoproterenol, aminophylline, and atropine (Sato and Sato 1984; 1988). Samples were collected at timed intervals under oil, and volume was measured in constant-bore calibration pipettes. CF patients (n = 14) did not sweat in response to this maneuver. The sweat-ductal potential difference (PD) was determined by measurement of the sweat-droplet PD under oil using a glass electrode filled with saturated KCl-agar and referenced to a subcutaneous electrode (Quinton and Bijman 1983; Bijman and Quinton 1984). The voltage for each subject is the average of 5–9 glands. The sweat-duct PD of 11 CF patients ranged from −46.3 to −80 mV. The nasal bioelectric parameters were measured as described elsewhere (Knowles et al. 1981, 1983*a*, 1983*b*). The mean nasal transepithelial PD and response to superfusion of drugs and ionic solutions were assessed via a Ringer-perfused exploring bridge referenced to a subcutaneous electrode. Inhibition with 10^{-4} M amiloride was calculated as percentage from baseline. The Cl^--diffusion PD (an index of Cl^- permeability) is the voltage change recorded during superfusion of a Cl^--free solution with 10^{-5} M isoproterenol in the presence of amiloride; a negative change indicates hyperpolarization.

Results

Family studies using allele-specific oligonucleotides to detect the normal and ΔF508 sequences revealed unusual results in three unrelated families (Lemna et al. 1990). Each family had an index parent whose DNA hybridized only to the ΔF508 oligonucleotide, which would imply that these apparently healthy parents were homozygous for the ΔF508 allele and were affected with CF. In addition, each index parent had one parent who was unexpected in that their DNA hybridized only with the normal oligonucleotide. Analysis of the amplified DNA from each index parent on a 10% acrylamide gel (Rommens et al. 1990) revealed heteroduplex formation, indicating that these individuals were heterozygous in some way and must carry a distinctive allele, which yielded amplified DNA that did not hybridize with either the normal or the ΔF508 oligonucleotides.

The major portion of exon 10 was amplified from genomic DNA from each index parent and the respective grandparent suspected of carrying the unusual chromosome. The C16B and C16D oligonucleotides (Kerem et al. 1989) were modified to add restriction-enzyme sites, and the products were cloned into M13 phage for DNA sequencing. As shown in figure 1, one mutant sequence included TGT at residue 508, causing a phenylalanine-to-cysteine substitution (F508C). A sec-

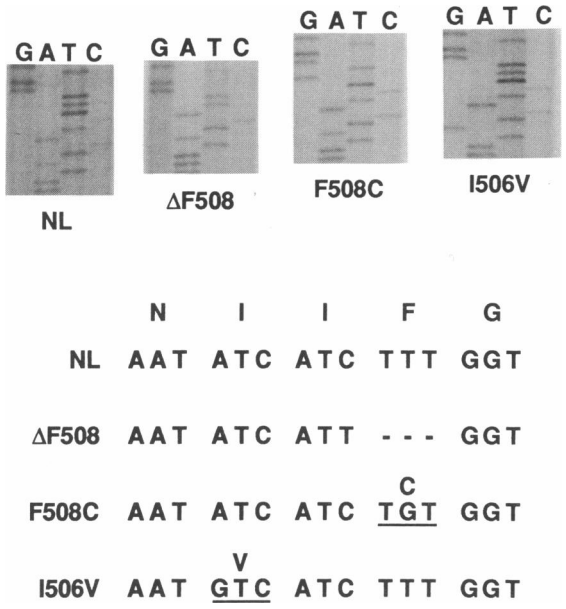

Figure 1 DNA sequence for two missense mutations. M13 clones from various individuals in the three families studies were sequenced, and examples of the normal and mutant sequences are shown. NL = normal; ΔF508, F508C, and I506V are mutant sequences described in the text. The nucleotide and amino acid sequence for positions 505–509 are depicted below.

ond mutant sequence included GTC at position 506, causing an isoleucine-to-valine substitution (I506V). The F508C mutation was found in two Caucasian families on a chromosome carrying the C haplotype for the XV-2c and KM-19 markers (Lemna et al. 1990). The other family was Hispanic, and the I506V mutation was present on a chromosome with the B haplotype. More extensive haplotypes for the mutant chromosomes were as follows: 1121122 for both chromosomes with the F508C mutation and 2112212 for the I506V chromosome. Haplotype markers in order are MetD/*Ban*I, MetD/*Taq*I, XV-2c/*Taq*I, KM-19/*Pst*I, E2.6/*Msp*I, J44/*Xba*I, and J3.11/*Msp*I (Kerem et al. 1989).

In order to confirm these mutations and study their inheritance within the three families, allele-specific oligonucleotides were prepared for each of the mutant sequences. Genomic DNA was amplified using the C16B and C16D primers, and hybridization was performed using four different allele-specific oligonucleotides (normal, ΔF508, F508C, and I506V) as shown in figure 2. For all three families (only two are shown in fig. 2), the child affected with CF was homozygous for the ΔF508 mutation. The index parent was a compound heterozygote carrying the ΔF508 allele and either the F508C or the I506V allele. In all cases, the index parent had a parent who was heterozygous for the ΔF508 allele and another parent who was heterozygous for either the F508C or the I506V allele. In one family depicted in figure 2, the compound heterozygous mother had two sisters with the same genotype. Using these allele-specific oligonucleotides, we rescreened approximately 440 normal and CF chromosomes and did not identify any other families carrying these alleles. This gave a frequency of approximately 2/440 apparently normal chromosomes carrying the F508C mutation and 1/440 apparently normal chromosomes carrying the I506V mutation. Neither of these mutations was ob-

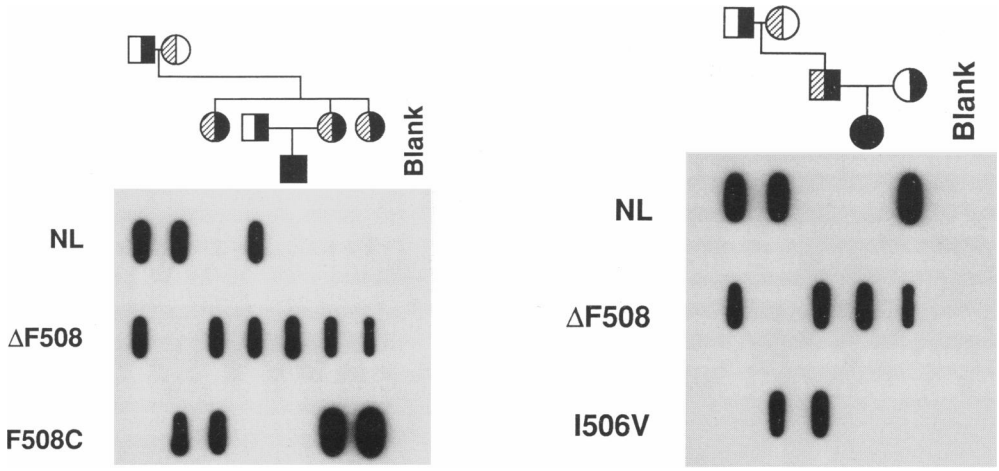

Figure 2 Family analysis with allele-specific oligonucleotides. Genomic DNA was amplified, applied to a slot-blot apparatus, and hybridized to each oligonucleotide. *Left,* Family W; *right,* family G. Solid gene symbols indicate individuals affected with CF. Solid half-symbols indicate the presence of one ΔF508 chromosome, and slashed half-symbols indicate the presence of an F508C or I506V chromosome.

Table 1

Physiological Studies on Three Compound Heterozygous Parents

Family[a]	Age	Sex	Forced Expiratory Volume in 1 s (% of predicted)	Sweat Cl^- (meq/liter)	Sweat Rate in Response to Isoproterenol (nl/min/gland)	Sweat-Duct PD (mV)	Nasal Bioelectric Parameters: Mean PD (mV)	Amiloride Inhibition (%)	Cl^--Diffusion PD (mV)
W	38	F	110	32,33	.8	−23.0	−28.3	49.5	−41.5
R	39	M	105	34,37	1.5	−23.6	−20.6	45.9	−47.0
G	36	M	88	19,22	1.2	−18.1	−15.1	56.9	−24.0
Normal range			80–120	<60	.2–3.0	−10 to −45	−10 to −40	<70	−15 to −60

Note.—The sweat Cl^- and forced expiratory volume in 1 s are standard clinical parameters. Other parameters are described in Material and Methods.

[a] The genotype for index parents from families W and R is ΔF508/F508C and that for the index parent from family G is ΔF508/I506V.

served on CF chromosomes, although an allele-specific oligonucleotide designed to hybridize to the I506V and the ΔF508 mutations on the same chromosome was not tested.

To determine whether these missense mutations were benign or caused disease, clinical evaluation and studies designed to detect expression of CF-specific respiratory or sweat-gland epithelial abnormalities were performed. The parent from family W had intermittent rhinitis, but no sinusitis or other pulmonary symptoms. The parent from family R reported episodes of "pneumonia" in 1972 and 1977, but had no chronic pulmonary complaints. The parent from family G smokes cigarettes (one pack per day for 20 years) and had an intermittent cough sometimes productive of mucopurulent secretions. Each of the three individuals had a normal physical examination, chest radiograph, and serum trypsinogen and vitamin E. Spirometry was normal in all three parents except for mild airflow obstruction at lower lung volumes in one parent (table 1, family G) compatible with his smoking history. Each of the three individuals exhibited normal sweat-gland and respiratory epithelial functional characteristics (table 1), including basal transepithelial bioelectric properties, response to cyclic AMP–dependent Cl^- secretagogues, and Cl^- permselectivity. Thus, these parents exhibited no evidence of the clinical disease or epithelial dysfunction which is seen with even "mild" CF (Stern et al. 1977; Evensen 1981; Knowles et al. 1989).

Discussion

Since the phenotype for the ΔF508 allele is known to be severe CF, the clinical and physiological data for the three compound heterozygote parents indicate that the F508C and I506V alleles must be benign. It is possible that these missense alleles could cause CF-like phenotypes in other genotypic combinations, although this seems unlikely. Since the entire gene was not sequenced, there is also a small possibility that the alleles carrying the F508C or I506V mutations could contain second mutations which balance a potential deleterious effect; this would leave open the possibility that the F508C or I506V mutations alone might be pathological. These F508C and I506V mutations might be found in the future on CF chromosomes carrying other deleterious mutations. This would be similar to the deleterious alleles at the α_1-antitrypsin locus which are found in various combinations with other benign amino acid substitutions within the gene (Cox 1989). This would likely occur by sequential independent mutation, although recombination and gene conversion are possible. At the present time, if the F508C or I506V alleles are found on CF chromosomes, it should be presumed that there is a deleterious mutation elsewhere in the gene until proven otherwise. The existence of these alleles further emphasizes how uncommon genetic variation can cause diagnostic pitfalls in molecular diagnosis. Variation in the DNA sequence at the site of allele-specific oligonucleotides or at the site of primers used for amplification (Fujimura et al. 1990) can lead to diagnostic errors.

The molecular, clinical, and physiological data in these three compound heterozygous parents indicate that the amino acid sequence in the F508 region of the CFTR need not be rigidly conserved in order to maintain adequate gene function. In particular, the F508C allele represents a nonconservative substitution, but

there is no detectable phenotypic effect. This information in combination with the knowledge that the ΔF508 (Kerem et al. 1989) causes the clinical syndrome of CF suggests that the length of this region of the CFTR protein is more important than the exact amino acid residues which are present. These missense mutations occur in a slightly less conserved portion of the nucleotide-binding fold between two more highly conserved motifs (Higgins et al. 1988). This type of in vivo mutational analysis is extremely valuable in assessing the structure-function relationships of the CF gene product.

Acknowledgments

We thank Gerald Feldman and Wanda Lemna for the initial identification of these families; Jewel Ward, Nancy Carpenter, and Raymond Lewandowski for referral and assistance in the evaluation of families W, R, and G, respectively; Nina Church for technical assistance; Lap-Chee Tsui, Francis Collins, and Jack Riordan for sharing helpful information and for critical reading of this manuscript; and Grace Watson for assistance in the preparation of the manuscript. This work was supported in part by NIH grants DK39617, HL34322, and RR00046 and by the Cystic Fibrosis Foundation.

References

Biggin MD, Gibson TJ, Hong GF (1983) Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc Natl Acad Sci USA 80:3963–3965

Bijman J, Quinton PM (1984) Influence of abnormal cystic fibrosis impermeability on sweating in cystic fibrosis. Am J Physiol 247:C3

Boat TF, Walsh MJ, Beaudet AL (1989) Cystic fibrosis. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The Metabolic Basis of Inherited Disease. McGraw-Hill, New York; pp. 2649–2680

Cox DW (1989) α_1-antitrypsin deficiency. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The Metabolic Basis of Inherited Disease. McGraw-Hill, New York, pp 2409–2437

Evensen SA (1981) A 69-year-old man with chronic obstructive pulmonary disease: pancreatic insufficiency and elevated sweat electrolytes. Acta Med Scand 209:141–143

Fujimura FK, Northrup H, Beaudet AL, O'Brien WE (1990) Genotyping errors with the polymerase chain reaction. N Engl J Med 322:61

Gibson LE, Cooke RE (1959) A test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilizing pilocarpine by iontophoresis. Pediatrics 23:545–549

Higgins CF, Gallagher MP, Mimmack ML, Pearce SR (1988) A family of closely related ATP-binding subunits from prokaryotic and eukaryotic cells. Bioessays 8:111–116

Kerem B-S, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, Buchwald M, Tsui L-C (1989) Identification of the cystic fibrosis gene: genetic analysis. Science 245:1073–1080

Knowles MR, Barnett RB, McConkie-Rosell A, Sawyer C, Kahler SG (1989) Mild cystic fibrosis in a consanguineous family. Ann Intern Med 110:599–605

Knowles MR, Gatzy JT, Boucher RC (1981) Increased bioelectric potential difference across respiratory epithelia in cystic fibrosis. N Engl J Med 305:1489–1495

——— (1983*a*) Modulation of nasal epithelial ion permeability in normal and cystic fibrosis subjects in vivo. Clin Res 31:858A

——— (1983*b*) Relative ion permeability of normal and cystic fibrosis nasal epithelium. J Clin Invest 71:1410–1417

Lemna WK, Feldman GL, Kerem B-S, Fernbach SD, Zevkovich EP, O'Brien WE, Riordan JR, et al (1990) Mutation analysis for heterozygote detection and prenatal diagnosis of cystic fibrosis. N Engl J Med 322:291–296

Quinton PM, Bijman J (1983) Higher bioelectric potentials due to decreased chloride absorption in the sweat glands of patients with cystic fibrosis. N Engl J Med 308:1185–1189

Riordan JR, Rommens JM, Kerem B-S, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, et al (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 245:1066–1073

Rommens J, Kerem BS, Greer W, Chang P, Tsui L-C, Ray P (1990) Rapid nonradioactive detection of the major cystic fibrosis mutation. Am J Hum Genet 46:395–396

Rommens JM, Iannuzzi MC, Kerem B-S, Drumm ML, Melmer G, Dean M, Rozmahel R, et al (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 245:1059–1065

Sato K, Sato F (1984) Defective beta adrenergic response of cystic fibrosis sweat glands *in vivo* and *in vitro*. J Clin Invest 73:1763–1771

——— (1988) Variable reduction in β-adrenergic sweat secretion in cystic fibrosis heterozygotes. J Lab Clin Med 111:511–518

Stern RC, Boat TF, Doershuk CF, Tucker AS, Miller RB, Matthews LW (1977) Cystic fibrosis diagnosed after age 13: twenty-five teenage and adult patients including three asymptomatic men. Ann Intern Med 87:188–191