A Common Mutation in the Surfactant Protein C Gene Associated with Lung Disease

H. Scott Cameron MD, Marco Somaschini MD, Paola Carrera MD, Aaron Hamvas MD, Jeffrey A. Whitsett MD, Susan E. Wert PhD, Gail Deutsch MD, and Lawrence M. Nogee MD

Department of Pediatrics, Division of Neonatology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21287, U.S.A. [H.S.C., L.M.N.], Unità di Patologia Neonatale, Ospedale Bolognini, Seriate (Bergamo), Italy [M.S.], Laboratorio di Biologia Molecolare Clinica, IRCCS H San Raffaele, Milano, Italy [P.C.], Edward Mallinckrodt Department of Pediatrics, Washington University School of Medicine, and St. Louis Children’s Hospital, St. Louis, Missouri 63110, U.S.A. [A.H.], Department of Pediatrics, Divisions of Neonatology and Pulmonary Biology, University of Cincinnati College of Medicine and Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, U.S.A. [J.A.W., S.E.W.], and Department of Pathology, University of Cincinnati College of Medicine and Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, U.S.A. [G.D.]

Supported in part by grants from the National Institutes of Health HL-54703 (L.M.N.), HL-56387 (J.A.W., S.E.W., L.M.N.), HL-65174 (A.H.), Forest Pharmaceuticals (H.S.C.) and the Eudowood Foundation.

Correspondence and reprint requests: Lawrence M. Nogee, M.D., Eudowood Neonatal Pulmonary Division, 600 North Wolfe Street/CMSC 6-104, Baltimore, MD 21287-3200, U.S.A.; e-mail: lnogee@jhmi.edu; phone: 410-955-5259; FAX: 410-955-0298

Key words: interstitial lung disease, sporadic, familial, infant, single nucleotide polymorphism

Running title: Common SP-C gene mutation
ABSTRACT

Objective: To determine the contribution of the surfactant protein C (SP-C) I73T mutation to lung disease.

Study Design: Genomic DNA was obtained from 116 children with interstitial lung disease (ILD) or chronic lung disease of unclear etiology and from 166 controls and was screened for the I73T mutation using an allele-specific polymerase chain reaction (PCR) assay.

Results: The I73T mutation was found on 7 of 232 SP-C alleles from seven unrelated children with ILD but was not found on 332 control SP-C alleles (p<0.01, Fisher’s exact test). The I73T mutation segregated with lung disease in one kindred with familial ILD. The I73T mutation was found in an asymptomatic parent from two different families with affected children consistent with variable penetrance, but it was not found in either asymptomatic parent of two other unrelated affected children consistent with a de novo mutation. Analysis of single nucleotide polymorphisms indicated diverse genetic backgrounds of the I73T alleles. Immunohistochemical analysis of lung tissue from an infant with the I73T mutation demonstrated normal staining patterns for proSP-B, SP-B, and proSP-C.

Conclusion: These findings support the hypothesis that the I73T mutation predisposes to or causes lung disease.
INTRODUCTION

Pulmonary surfactant is the mixture of lipids and specific proteins needed to reduce surface tension at the air-liquid interface and prevent end-expiratory atelectasis. Lack of surfactant due to immaturity is the principal cause of the respiratory distress syndrome (RDS) observed in infants born prematurely (1). RDS is effectively treated with exogenous surfactants that contain the hydrophobic surfactant proteins SP-B and SP-C (2). The importance of these two proteins in normal lung function is highlighted by lung diseases due to inherited abnormalities in their genes (3, 4). Loss of function mutations on both alleles of the SP-B gene result in hereditary SP-B deficiency, an autosomal recessive cause of fatal neonatal respiratory disease (5). SP-C deficiency and mutations in the SP-C gene have been associated with interstitial lung disease (6).

Three mutations in the SP-C gene have been reported in individuals with familial and sporadic interstitial lung disease (7-9). These SP-C gene mutations resulted in the production of an abnormal SP-C precursor protein (proSP-C) and were present on only one allele in affected individuals. These findings are consistent with the observed autosomal dominant inheritance pattern and suggested that the mutated proSP-C may have interfered with processing of wild-type proSP-C. It is unclear whether lung disease was caused by abnormal proSP-C blocking production of the active peptide or inducing toxicity in alveolar type II (AEP II) cells. The relative contributions of these SP-C gene mutations to different types of interstitial lung disease as well as the range of phenotypes associated with these SP-C gene mutations are unknown. The overall incidence of lung disease related to SP-C gene mutations is also unknown. The finding
of a common mutation has proven useful both for diagnosis and providing an estimate of disease frequency in other genetic diseases (10-12).

In the present study, we have identified a novel SP-C gene mutation in multiple unrelated infants with lung disease. The finding of the same mutation in unrelated individuals with lung disease could indicate that it is a disease-causing mutation or that it represents a rare but benign polymorphism. To address this question, we have compared the frequency of this mutation in a cohort of patients with ILD and a control group without lung disease, examined the segregation pattern of the mutation with disease in several families, and compared the genetic backgrounds of alleles containing the mutation.
METHODS

Patients and DNA
Patient samples were obtained as part of a collaborative study involving the Johns Hopkins Children’s Center, St. Louis Children’s Hospital, and Cincinnati Children’s Hospital to evaluate infants with lung disease of unknown cause for mutations in the surfactant protein genes. Samples were also recruited from physicians from other medical centers familiar with our work who were interested in having their patients evaluated for potential SP-B and SP-C gene abnormalities. The institutional review boards of the participating institutions approved the protocols for these evaluations, and written informed consent for genetic testing was obtained from the families. The experimental group consisted of 116 children with interstitial lung disease or chronic lung disease of unclear etiology. Chronic lung disease was defined as persistent respiratory symptoms and/or a supplemental oxygen requirement with chest radiographic changes of diffuse granular infiltrates or increased interstitial markings. The control group consisted of 166 adults without a history of lung disease. Genomic DNA was prepared from blood leukocytes with the use of a commercially available kit (Gentra Systems). Experimental samples were screened for all published SP-C gene mutations (7-9). Control samples were analyzed anonymously.

Mutation Analysis
To efficiently screen for the I73T mutation, we developed a PCR based assay using a primer specific for the I73T allele. This primer contained the thymine to cytosine transition at its 3’ end, which only allowed for amplification of alleles containing the I73T mutation under
appropriate conditions. The SP-C I73T allele-specific primer (5’GGTTCTGGAGATGAGCAC3’) was combined with an antisense primer (5’CTGGAAAGTTTGAGACTTTTC3’) to generate a 467 bp product. An internal positive control for amplification was included in each reaction using primers specific for a portion of the SP-B gene sequence (sense primer 5’CTTGCCAAGTGAGGTCC3’; antisense primer 5’CAGCCTTCTGCCTTGAACG3’) to generate a 606 bp product. PCR conditions were optimized to reproducibly generate this 467 bp product in all individuals who carried the I73T mutation. The PCR conditions were: initial denaturation 95 °C for 4 ½ minutes; denaturation 95 °C for 30 seconds, annealing 60 °C for 30 seconds, extension 72 °C for 45 seconds for 30 cycles; and final extension 72 °C for 10 minutes. The PCR products were analyzed by agarose gel electrophoresis. Assays positive for the I73T mutation were confirmed by restriction analysis with Bsp1286I, which cuts at the recognition site created by I73T, and direct DNA sequencing through the institutional core DNA sequencing facility.

**Single Nucleotide Polymorphism (SNP) Analysis**

SNP analysis of the I73T alleles was performed using the I73T allele-specific primer in both the sense (above) and antisense orientations coupled with primers from the SP-C 3’ untranslated (5’TGCAAGGAAGACGAGCAGACC3’) and intron 1 (5’AGCATAGCACCTGCAGCAAG3’) regions, respectively. Using the same PCR conditions mentioned above, an initial PCR reaction with the above allele-specific primer sets was followed by a nested PCR reaction using internal primers to generate PCR products to be sequenced. All allele-specific PCR products were analyzed by agarose gel electrophoresis, purified with the use of a commercially available kit (Qiagen), and sequenced.
**Immunohistochemical Analysis**

Immunohistochemical analysis was performed on formalin-fixed, paraffin-embedded tissue as previously described (5, 13). The production and characterization of polyclonal antiserum against surfactant protein B precursor protein, surfactant protein B, and surfactant protein C precursor protein have been described previously (14-17).
RESULTS

The I73T mutation was first identified by DNA sequence analysis of the SP-C gene in a 33 weeks gestation female infant who initially presented with symptoms consistent with RDS but failed to wean off supplemental oxygen by 2 years of age. Evaluation was prompted when her mother was diagnosed with restrictive pneumonitis post-partum, and additional family history revealed an extensive familial lung disease kindred (Figure 1). A heterozygous thymine to cytosine transition was identified at nucleotide 218 corresponding to the SP-C cDNA sequence that resulted in the substitution of threonine for isoleucine (ATT > ACT) at codon 73 (I73T). Segregation analysis demonstrated the presence of the mutation in patients II:3, II:4, III:1, III:2, III:3, and III:4, all with lung disease (Figure 1). The same mutation was also found in an unrelated six-year-old girl with pulmonary fibrosis who was being evaluated for lung transplantation.

After finding the same mutation in two unrelated individuals with lung disease, we screened 116 children with interstitial lung disease or chronic lung disease of unclear etiology and 166 adults without a history of lung disease for the I73T mutation using an allele-specific PCR assay. The I73T mutation was found on 7 of 232 SP-C alleles from patients with lung disease and on 0 of the 332 SP-C alleles from control samples (p<0.01, Fisher’s exact test). The presence of the I73T mutation was confirmed by restriction digests and direct DNA sequencing (data not shown). One patient with the I73T mutation had an additional SP-C gene mutation, L110R, identified on her other allele. Nine additional SP-C gene mutations were identified in the experimental group (). Details of these mutations will be published in a separate report.
Clinical information for each of the seven patients with the I73T mutation is summarized in Table 1. All seven patients were female infants with gestational ages ranging from 33 weeks to term. One patient was African-American and the remaining six were Caucasian. The age of onset of respiratory symptoms ranged from birth to 24 months of age. Five of the seven children underwent lung biopsy with histopathology interpreted as chronic pneumonitis of infancy (CPI) or desquamative interstitial pneumonitis (DIP). All patients are currently alive. Six patients have a persistent supplemental oxygen requirement, and one patient underwent lung transplantation at 5½ years of age. Two patients are awaiting lung transplantation.

DNA was obtained from one or both parents of 5 of the 7 children with the I73T mutation (Table 1). Figure 2 shows the limited pedigree from three of these families. The proband (Patient 1) is included for comparison. In two families (Patients 3 and 4), one parent was found to be a carrier of the I73T mutation. Additionally, the mother of patient 4 was found to be a carrier of a different SP-C gene mutation, L110R. All of these individuals are currently asymptomatic but have not been formally evaluated for pulmonary disease. In two other families (Patients 5 and 7), neither asymptomatic parent carried the mutation, which is consistent with de novo mutations resulting in sporadic disease.

To evaluate surfactant protein expression in an individual with the I73T mutation, immunohistochemical analysis was performed on lung tissue obtained from lung biopsy. Haematoxylin and eosin staining is depicted in Figure 3, Panel A. Strong staining for proSP-C was present in the alveolar epithelium as shown in Figure 3, Panel B. ProSP-B and SP-B were also present as demonstrated by strong staining with their respective antibodies (data not shown).
Histopathologic features include diffuse alveolar septal thickening, prominent AEP II cell hyperplasia, and accumulation of macrophages in the alveolar lumen.

The finding of a common mutation could reflect a shared ancestral origin or result from recurrent \textit{de novo} events. We next examined the genetic background of the different \textit{SP-C} alleles on which the I73T mutation was found. The I73T allele-specific primer and its reverse complement were used to generate longer PCR products in both the 5’ and 3’ directions such that sixteen SNPs on the I73T allele from each of the seven children were determined. We found 5 distinct \textit{SP-C} alleles that contained the I73T mutation (Figure 4). Patients 1 and 2 were identical for one combination of SNPs, while patients 3 and 5 shared a different pattern of SNPs. Each of the remaining three patients had a different SNP combination. These results suggest that the I73T allele has occurred in at least 5 different genetic backgrounds.
DISCUSSION

We have identified a novel SP-C gene mutation, I73T, in multiple unrelated infants with lung disease. While adult family members with the mutation who do not currently have lung disease were identified, incomplete penetrance has been recognized with another SP-C gene mutation with onset of lung disease in some individuals as late as the 6th decade (8). The segregation of the mutation with disease in the one extended kindred, the absence of the mutation in healthy controls, and the *de novo* occurrence of the mutation in association with disease suggest that the I73T mutation either causes or predisposes to lung disease.

It is possible that this mutation represents a rare but benign polymorphism. The corresponding codon of the SP-C mRNA has not been strictly conserved during evolution, and threonine and isoleucine are both neutral amino acids. However, the substitution of the polar amino acid threonine for the non-polar amino acid isoleucine in other proteins such as alpha-spectrin and fibrillin-1 has been associated with other autosomal dominant disorders (elliptocytosis and neonatal Marfan Syndrome, respectively) (18,19). This similarity supports the notion that I73T is a disease-causing mutation. Conversely, the substitution of threonine for isoleucine at codon 131 of the SP-B gene due to a SNP at nucleotide 1580 (C/T) does not result in SP-B deficiency or clinical disease. However, the SP-B codon 131 Thr allele is over-represented in patients who develop RDS and acute respiratory distress syndrome (ARDS), which suggests that with certain environmental exposures it may predispose individuals to lung disease (…,20). The phenotypic variability observed in patients with the I73T mutation indicates a role for environmental and/or
genetic modifiers in contributing to the pathogenesis of lung disease in patients with this mutation.

The I73T mutation shares several features with previously characterized SP-C gene mutations. All published SP-C gene mutations are located in the carboxy-terminal domain of proSP-C, as is the I73T mutation in exon 3 (7-9). Deletions in this domain have been shown to result in incomplete processing of the mutant proSP-C, trapping of the proprotein in the endoplasmic reticulum, and rapid degradation via a proteasome-dependent pathway (21-23). All published SP-C gene mutations were only present on one allele in affected individuals, which is consistent with the observed autosomal dominant inheritance pattern (7-9). The I73T mutation was also only found on one allele in affected individuals. This inheritance pattern suggests that the abnormal proSP-C resulting from the mutations has a dominant negative effect on SP-C function and/or metabolism. Taken together, these similarities further suggest that the I73T mutation is associated with lung disease.

While immunohistochemical analysis in one of the seven patients demonstrated normal staining patterns for proSP-B, SP-B, and proSP-C in AEP II cells, there was diffuse alveolar septal thickening, prominent AEP II cell hyperplasia, and accumulation of macrophages in the alveolar lumen. These findings are consistent with what has been observed with previously characterized SP-C gene mutations. These findings are also consistent with one of the proposed pathogenesis mechanisms of SP-C gene mutations, that is, abnormally folded proSP-C forms oligomers with both mutant and wild-type proSP-C resulting in secondary cellular injury and subsequent inflammation in AEP II cells. Still, a lack of mature SP-C in secreted surfactant may also be a
contributing factor to the pathogenesis of SP-C gene mutations as SP-C knockout mice develop abnormal lung function in adulthood (24). Bronchoalveolar lavage fluid from a patient with the I73T mutation was not available to examine surfactant protein levels and investigate this possibility.

All seven infants with ILD who carried the I73T mutation were female. Parents and extended family who carried the I73T mutation and were symptomatic demonstrated an equal gender distribution. This finding was also seen with the exon 5 + 128 T>A mutation where an extended pedigree included fourteen patients with ILD of which three had been diagnosed as infants, two female and one male (8). Gender specific differences in the risk for lung disease in children has also been observed for a SP-B gene polymorphism and neonatal RDS (). In addition to other genetic and environmental modifiers, incomplete penetrance of the I73T mutation may be gender-specific when lung disease presents early in life.

One patient in this series had a second mutation, L110R, identified on her other SP-C allele, which results in a non-conservative neutral to basic amino acid substitution. It is possible that this mutation is responsible for her disease, or it could also represent a rare but benign polymorphism. Neither parent, each of whom carries one of the mutations, is known to have lung disease at this time. However, this child has relatively severe lung disease compared to the other patients in this series and is currently being evaluated for lung transplantation. It is thus possible that both alleles have contributed to the development of her lung disease.
The asymptomatic parents of two unrelated children with apparent sporadic disease did not carry the mutation, which is consistent with *de novo* mutations. Because of this apparent sporadic disease, we examined the genetic background of the different SP-C alleles on which the I73T mutation was found. The finding of the I73T mutation on diverse genetic backgrounds suggests that the mechanism for this common mutation resulted largely from recurrent *de novo* events. We speculate that this nucleotide position may represent a region of the genome that is particularly susceptible to mutation ("hot spot"). While non-paternity and alternative mechanisms, such as germ-line mosaicism, have not been excluded, the pattern of SNPs for several of the infants was consistent with maternal and paternal contribution.

In summary, we have identified a SP-C gene mutation in individuals with ILD and developed a method to rapidly screen for this mutation. The presence of this mutation on diverse SP-C alleles by SNP analysis suggests that this site may represent a "hot spot" for recurrent *de novo* events. The association of the I73T allele with lung disease, its absence in healthy controls, and its *de novo* occurrence in association with disease support the hypothesis that the I73T mutation either causes or predisposes to lung disease and is not simply a benign polymorphism. *In vitro* studies of the effects of the I73T mutation in cell culture and ultimately animal models will be needed to confirm this hypothesis. Our studies support evaluating individuals with ILD for this common SP-C gene mutation. The finding of a relatively common SP-C gene mutation associated with ILD may be useful in larger population-based studies to help estimate the disease frequency of ILD secondary to SP-C gene mutations.
ACKNOWLEDGEMENTS

We would like to thank Dr. Hal Dietz and Dr. Gary Cutting for the generous contribution of anonymous adult genomic DNA specimens used as controls.

ABBREVIATIONS

SP-C  surfactant protein C
SP-B  surfactant protein B
ILD  interstitial lung disease
DNA  deoxyribonucleic acid
mRNA  messenger ribonucleic acid
PCR  polymerase chain reaction
RDS  respiratory distress syndrome
ARDS  acute respiratory distress syndrome
AEP II  alveolar type II
SNP  single nucleotide polymorphism
CPI  chronic pneumonitis of infancy
DIP  desquamative interstitial pneumonitis
REFERENCES


dysfunction with a deletion of surfactant protein C gene. *Am J Respir Cell Mol Biol.*


high-throughput DNA extraction and molecular amplification from Guthrie cards.

analysis of CFTR mutations--correlation with incidence data and application to

13. Nogee LM, de Mello DE, Dehner LP, Colten HR. Brief report: deficiency of


15. Baatz JE, Elledge B, Whitsett JA. Surfactant protein SP-B induces ordering at the

16. Lin S, Akinbi HT, Breslin JS, Weaver TE. Structural requirements for targeting of
surfactant protein B (SP-B) to secretory granules in vitro and in vivo. *J Biol Chem.*


25. Glasser SW, Korfhagen TR, Perme CM, Pilot-Matias TJ, Kister SE, Whitsett JA.

Two SP-C genes encoding human pulmonary surfactant proteolipid. *J Biol Chem.*

FIGURE LEGENDS

Figure 1. Pedigree for the index patient.
Generation (I-IV): identifier number and clinical information for those with lung disease is shown below their symbol. Circle = female, square = male, open symbol = unaffected by lung disease, closed symbol = affected by lung disease, asterisk = I73T carrier, and slash through symbol = deceased.

Figure 2. Segregation of the I73T allele.
Circle = female, square = male, open symbol = unaffected by lung disease, closed symbol = affected by lung disease, and asterisk = I73T carrier. Allele-specific PCR results are shown below each corresponding pedigree symbol. The negative control is denoted by H20. The positive control is the 606 bp fragment which corresponds to part of the SP-B gene amplified in all patients. The experimental is the 467 bp fragment present only in patients with the I73T mutation.

Figure 3. Immunohistochemical Staining for Surfactant Proteins.
H & E staining from a patient with the I73T mutation (Panel A). ProSP-C staining is shown in Panels B [size bar and original magnification (X 150) are the same for each panel].

Figure 4. SP-C Single Nucleotide Polymorphisms.
The pattern of single nucleotide polymorphisms (SNPs) for patients 1 through 7 is shown. The SP-C gene is illustrated at the top with the location of each SNP denoted by an arrow. The
location of the I73T mutation is also shown. The specific nucleotides at each genomic position from 5’ to 3’ are: rs8192325 (1=A, 2=G); g1335 (1=T, 2=C); rs8192327 (1=G); rs2070684 (1=A, 2=C); rs8192341 (1=A, 2=C); rs8192330 (1=G); rs2070685 (1=C, 2=T); rs2070686 (1=C); rs2070687 (1=C); rs1124 (1=A, 2=G); g3059 (1=insert A, 2=no insert A); rs1126948 (1=C); g3036 (1=A); rs7592 (1=A, 2=G); rs1126931 (1=C, 2=T); rs1139547 (1=A, 2=G) where G = guanine, A = adenine, T = thymine, and C = cytosine. The polymorphisms are labeled either by their entry code in the public data base of SNPs (http://www.ncbi.nlm.nih.gov/SNP/) or by their positions based on the most recent SP-C genomic sequence (25).