Effect of genotype on the levels of surfactant protein A mRNA and on the SP-A2 splice variants in adult humans

Anne M. KARINCH*, Daphne E. DEMELLO† and Joanna FLOROS*‡§

Departments of *Cellular and Molecular Physiology and ‡Paediatrics, The Pennsylvania State University, College of Medicine, Hershey, PA 17033, and †Department of Pathology and Laboratory Medicine, Cardinal Glennon Children's Hospital and St. Louis University Health Services Center, St. Louis, MO 63104, U.S.A.

Human pulmonary surfactant protein A (SP-A) is encoded by two genes, SP-A1 and SP-A2, that exhibit coding sequence (allelic) and 5′ splicing variability. In this report we determine the effect of the genetic variability within the SP-A1 and SP-A2 genes on the level of SP-A mRNAs and on the SP-A2 splicing variants in different individuals. We analysed mRNA specimens from 23 unrelated adults using genotype analysis, Northern analysis and primer extension, and made the following observations. (1) The level of SP-A mRNA varies among individuals (coefficient of variation = 0.49). One SP-A genotype ($6A^{2}6A^{2}1A^{0}1A^{0}$) appears to be associated with a low to moderate level of SP-A mRNA. (2) The SP-A1/SP-A2 mRNA ratio varies among individuals, from 0.94 (lowest) to 6.80 (highest) within the study population. One genotype appears to be associated with a moderate to high SP-A1/SP-A2 mRNA ratio and another with a low to moderate ratio. (3) There is no correlation between the level of SP-A mRNA and the SP-A1/SP-A2 mRNA ratio. (4) Variability in the ratio of the major SP-A2 splice variants among individuals results from nucleotide differences in the splice-recognition sequence of specific SP-A2 alleles. The SP-A mRNA levels, the SP-A1/SP-A2 mRNA ratio, and the ratio of the major SP-A2 splice variants have a genetic basis in that they vary depending upon the specific SP-A alleles present.

INTRODUCTION

Surfactant protein A (SP-A) is the major non-serum pulmonary surfactant-associated protein. Functional SP-A is a large multimeric (18 SP-A monomers) complex [1,2] involved in varied aspects of surfactant biology [3–5]. SP-A expression is tissue specific and is regulated by developmental stage and a number of hormones (reviewed in [6]). It is increasingly evident that human SP-A is also complex at the molecular level (reviewed in [7]). In humans, SP-A is encoded by two functional genes, SP-A1 and SP-A2. The two genes have very similar genomic structures and encode proteins that differ in eight amino acids [8,9], four of which lie within the collagen-like region of the protein that is responsible for the formation of SP-A trimers, the first step in the assembly of the functional complex. It is not known whether SP-A1 and SP-A2 differ in function, though it has been proposed on theoretical grounds that both proteins are required, in the ratio two SP-A1 to one SP-A2, for formation of the multimeric SP-A complex [10].

The genes exhibit variability on three levels. (a) Alternative splicing of 5′-untranslated (UT) exons results in major transcripts of both genes that differ in their 5′ UT region [11,12]. (b) Sequence variability within the coding region of both genes results in a number of allelic variants for each gene ([8,9,11–15]; the present paper). (c) The 3′ UT region of each gene contains considerable sequence variability [13,16,17]. It is currently not known whether some or all of this genetic variability has physiological significance, though it has been established that the major 5′ UT region splice variants of both SP-A1 and SP-A2 are translated *in vivo* [18].

The level of SP-A mRNA varies considerably among adult individuals [19]. The variability, as reflected by the coefficient of variation (CV; standard deviation/mean) of 0.53, was high compared with that observed for the hydrophobic surfactant protein SP-C (CV = 0.27) [19]. At the time that this initial observation, i.e. that the level of SP-A mRNA varies among adult humans, was made [19], the molecular complexity of SP-A was not fully appreciated. A recent study has shown that SP-C, in contrast with SP-A, is highly conserved among humans [20]. The apparent coincidence of molecular complexity with varied mRNA level (SP-A), and molecular conservation with relatively constant mRNA level (SP-C), is intriguing, and suggested to us that the observed variation in SP-A mRNA may be a direct result of its genetic heterogeneity. Consequently, we designed the studies described in the present report to begin to search systematically for links between SP-A mRNA expression and specific forms of SP-A genetic variability.

EXPERIMENTAL

Tissue acquisition

The human lung tissue used for these studies was discarded surgical tissue from patients who underwent lung resections primarily for tumours. The tissues used were considered 'normal' on gross inspection by a pathologist. In addition, sections of frozen and paraffin-embedded tissue from fifteen of the specimens (marked with triangles in Table 2) were stained with haematoxylin and eosin and were examined histologically by one of us (D. E. deM.). These tissue sections contained no major pathology and were from comparable levels of the lung, as determined from the size of small airways.

RNA isolation

Frozen human lung tissue was powdered and homogenized in buffer containing 4 M guanidinium thiocyanate. The DNA was sheared by passing repeatedly through a needle, and RNA was

Abbreviations used: SP-A, surfactant protein A; UT, untranslated; CV, coefficient of variation; RT, reverse transcription; BAL, bronchoalveolar lavage.

§ To whom correspondence should be addressed.

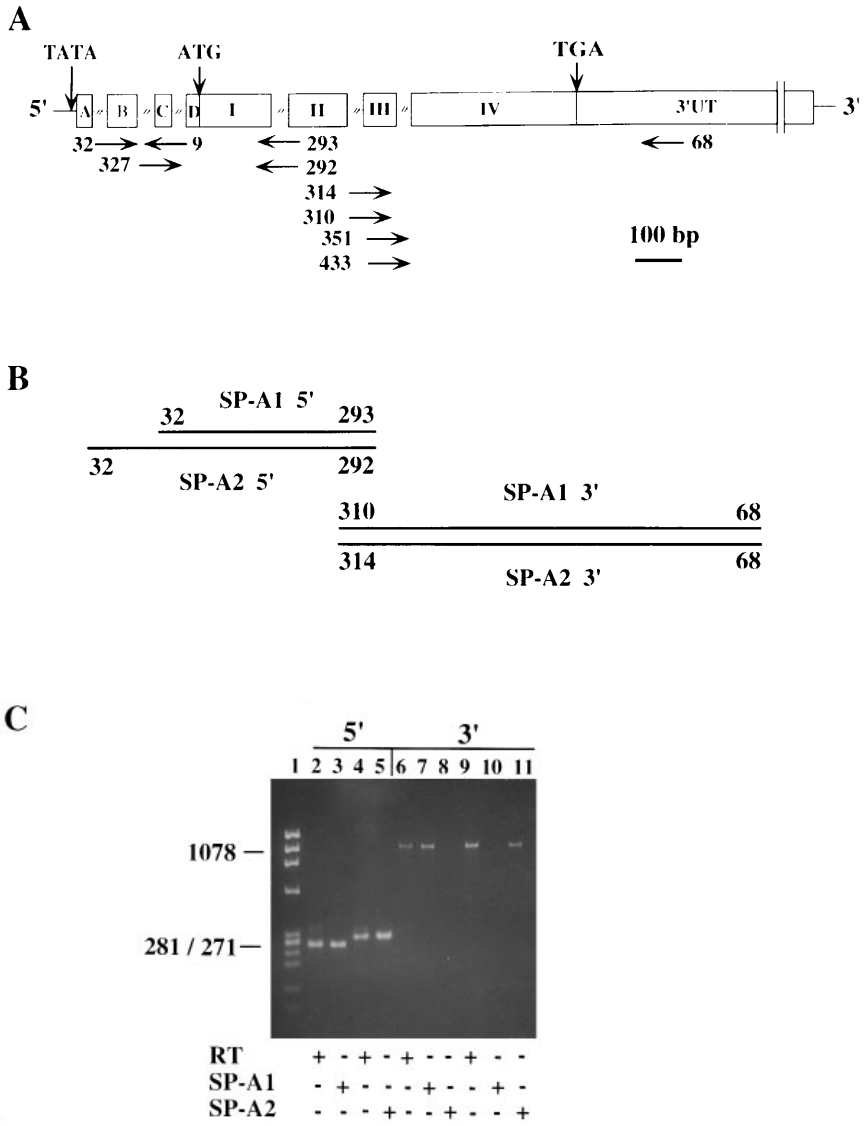

Figure 1 Specific amplification of SP-A cDNA

(**A**) The genomic structure of the SP-A gene is shown. The open boxes represent the exons, drawn to scale: 5′ UT exons, A through D; coding exons, I through IV; 3′ UT. The TATA box and translation start and stop sites are indicated by arrows above the boxes. Below the boxes are the locations of oligonucleotides used in the study; their orientation is indicated by arrows. (**B**) SP-A1 and SP-A2 cDNAs were divided, by PCR amplification, into two gene-specific fragments, as shown. The fragments were applied to slot blots for genotyping. (**C**) The specificity of amplification is shown. The 5′ (lanes 2–5) and 3′ (lanes 6–11) fragments were amplified as described in the Experimental section. The cDNA template included in each amplification mixture is indicated under each lane of the figure as follows: RT, RT reaction products; for the SP-A1 gene control, the allele $6A^3$ cDNA was used; for the SP-A2 gene control, the allele $1A^0$ cDNA was used. The products (RT or cDNA) in lanes 2, 3 and 6–8 were amplified with SP-A1-specific oligonucleotides, and the products in lanes 4, 5 and 9–11 were amplified with SP-A2-specific oligonucleotides. The size difference between the 5′ fragments from SP-A1 and SP-A2 can be seen (lanes 2–5). The faint higher band in the RT lanes (2 and 4) is the splice variant AB′D′, produced from both genes [11]. The specificity of the 3′ fragments is demonstrated by the amplification of the specific control cDNAs in lanes 7 and 11 and by the lack of product in lanes 8 and 10. Specificity of the 3′ fragments was further verified by gene-specific hybridization with oligonucleotides 351 and 433 (results not shown). Lane 1, Φ174 *Hae*III markers.

isolated by centrifugation through a caesium chloride cushion [21].

Northern analysis

Total lung RNA (10 μg) was separated on a formaldehyde gel and transferred to nylon membrane, according to the vendor's specifications (DuPont–NEN Life Science Products, Boston, MA, U.S.A.). Membranes were hybridized with SP-A cDNA and actin-specific oligonucleotide 182. Hybridization was quantified by scanning laser densitometry.

Reverse transcription (RT)-PCR for genotyping

RT was carried out as described in a total volume of 25 μl using mouse mammary leukaemia virus reverse transcriptase (Gibco/BRL, Gaithersburg, MD, U.S.A.). For RT of total lung RNA, 1 μg RNA was incubated with 15 ng of primer (oligonucleotide 68) at 70 °C for 10 min. The reaction was incubated at room temperature for 15 min with buffer (provided by the vendor of the reverse transcriptase)/0.5 mM each dNTP/10 mM DTT/0.5 μl of RNase block (Stratagene, La Jolla, CA, U.S.A.). Mouse mammary leukaemia virus reverse transcriptase (0.75 μl, 150 units) was added, and the reaction mixture was incubated at

Table 1 Oligonucleotides

The nucleotide positions for the SP-A1 gene are those of White and co-workers [8], and those for the SP-A2 gene are those of Katyal and co-workers [9].

Number of oligonucleotide	5′-to-3′ sequence	Orientation	Nucleotide position	Gene specificity
9	GGCACAGCCACATGGCTCTG	Antisense	1039–1058	SP-A1, SP-A2
32	CTGGAGGCTCTGTGTGTGGG	Sense	181–200	SP-A1, SP-A2
68	TGCCACAGAGACCTCAGAGT	Antisense	3845–3864	SP-A1, SP-A2
292	CCATTATTCCCAGGAGGACATGGTG	Antisense	1575–1599	SP-A2
293	CCATCATTTCCAGGAGGACATGGCA	Antisense	1555–1579	SP-A1
310	TGCCATGTCCTCCTGGAAATG	Sense	1555–1575	SP-A1
314	CACCATGTCCTCCTGGGAATA	Sense	1575–1595	SP-A2
327	ATCACTGACTGTGAGAGGGT	Sense	472–491	SP-A2
351	GGTATCCCTGGAGAGTGTGG	Sense	1595–1615	SP-A1
433	GGTGTCCCTGGAGAGCGTGG	Sense	1616–1635	SP-A2

46 °C for 1 h. The reaction was terminated by heating at 95 °C for 5 min. For genotyping samples, gene-specific amplification of SP-A cDNA was achieved by dividing the cDNA into 5′ and 3′ segments and amplifying each segment using SP-A1- and SP-A2-specific primers. For the 5′ segment, oligonucleotides 32 and 293 (SP-A1-specific) or 292 (SP-A2-specific) were used; for the 3′ segment, oligonucleotides 68 and 310 (SP-A1) or 314 (SP-A2) were used for amplification. The position and orientation of oligonucleotides are shown in Figure 1A, and the 5′ and 3′ fragments are shown in Figure 1B. For the 5′ segment, 2.5 μl of the RT reaction mixture was used as template in a total volume of 50 μl. For the 3′ segment, 5 μl of the RT reaction mixture was used in 100 μl. The PCR mixture contained buffer supplied by the polymerase vendor, 50 ng of each of the appropriate oligonucleotide pairs, 75 mM each dNTP, 1.5 mM $MgCl_2$ and 1 unit of AmpliTaq polymerase. Cycling conditions were 94 °C for 20 s, 53 °C for 20 s and 72 °C for 30 s, for 30 cycles (5′ segment) or 40 cycles (3′ segment), plus a final 72 ° extension for 10 min. The sizes of the amplified 5′ fragments of SP-A1 cDNA and SPA2 cDNA differ by 30 nucleotides as a result of alternative splicing of exon B. Therefore, the specificity of amplification of the 5′ fragment was verified by checking the size of the products on an agarose gel (Figure 1C). The specificity of amplification of the 3′ fragment was monitored by parallel amplification of characterized SP-A1 and SP-A2 cDNA clones (Figure 1C) and/or hybridization with a gene-specific oligonucleotide probe (oligonucleotides 351 and 433).

Primer extension

Primer extension was carried out as previously described [11], using 25 μg of lung total RNA and 150000 c.p.m. of ^{32}P-labelled oligonucleotide 9. Extension products were separated on a 6% polyacrylamide sequencing gel, and the sizes of the products were determined from a sequence ladder run on the same gel. Extension products of interest were quantified by densitometry of an autoradiograph of the gel.

Slot-blot/genotype analysis

The SP-A1 alleles determined by sequence information and characterized experimentally are 6A, 6A^2, 6A^3, and 6A^4. Similarly, the SP-A2 alleles are 1A, 1A^0, 1A^1, 1A^2 and 1A^3 [14]. RT-PCR products were applied to slot blots and were hybridized with 5′-end-labelled SP-A1 and SP-A2 allele-specific oligonucleotides as described by Floros and colleagues [14]. Hybridization and wash temperatures differed for the various oligonucleotides. The specificity of hybridization was monitored by inclusion on the slot blots of cDNAs of known SP-A1 and SP-A2 alleles [11].

Sequencing of splice recognition sequence

SP-A2 genomic DNA of the 1A, 1A^1, 1A^2 alleles of interest was amplified with oligonucleotides 327 and 68, and the 3.4 kb product was cloned into the pGemT vector (Promega, Madison, WI, U.S.A.). The clones were cycle-sequenced using the fmol sequencing system (Promega).

Oligonucleotides

Detailed information about the oligonucleotides used in the present study is presented in Table 1. Oligonucleotides 9, 32 and 68 were used for RT-PCR. Oligonucleotide 9 was used for the primer extension experiment. Oligonucleotides 351 and 433 were used to verify the specificity of the gene-specific amplification of SP-A 3′ fragment for genotyping. Gene- and allele-specific oligonucleotides used for genotyping are described in detail elsewhere [14].

Statistical analysis

Student's t test was used for comparison of total SP-A mRNA and the SP-A1/SP-A2 mRNA ratio between different groups. Haplotype analysis was carried out using the EH program [22].

RESULTS

As described in the Introduction, the SP-A genes exhibit a considerable degree of variability. In previous studies we have described four alleles for SP-A1 and five for SP-A2, based on nucleotide differences within the coding exons [11,14]; we observed that the ratio of the major SP-A2 5′ UT region splice variants varied among individuals [11,18]; we also reported that the level of SP-A mRNA varies among adult individuals [19]. The studies in which these observations were made involved a relatively small number of individuals: three [11], five [18] and eleven [19]. Therefore, we designed a series of experiments using a larger study population, from which we hoped to understand the basis of the variability in the amount of SP-A mRNA and the SP-A2 splicing variability among humans. In the present study we isolated total RNA from lung tissue of 23 unrelated adult individuals and subjected these RNAs to a number of different analyses. The SP-A genotypes of the study population were also

Table 2 Total SP-A mRNA and SP-A genotype

Relative levels of total SP-A mRNA (in arbitrary units) were determined from Northern-blot analysis for 23 adults. The SP-A genotype of these individuals was determined from gene- and allele-specific hybridization as described in the Experimental section. The individuals have been listed in order of increasing total SP-A mRNA, with their genotypes also listed. The genotype $6A^2 6A^2 1A^0 1A^0$ (†) appears to be associated with low-to-moderate levels of SP-A mRNA. Lung tissue from individuals marked with a triangle was subjected to microscopic analysis. The genotype $6A^2 6A^3 1A^0 1A^1$ (*) is discussed in the text.

Individual	Total SP-A mRNA (arbitrary units)	Genotype
1810	0.49	$6A^2\ 6A^3\ 1A^0\ 1A^{1*}$
1801▲	0.55	$6A^2\ 6A^2\ 1A^0\ 1A^{0}$†
1792	0.67	$6A^2\ 6A^3\ 1A^0\ 1A^{1*}$
1787▲	0.67	$6A^2\ 6A^3\ 1A^0\ 1A^2$
1806▲	0.73	$6A^2\ 6A^2\ 1A^0\ 1A^{0}$†
1807▲	0.83	$6A\ 6A^2\ 1A\ 1A^0$
1800▲	0.90	$6A^2\ 6A^2\ 1A^0\ 1A^{0}$†
1802	1.04	$6A^2\ 6A^2\ 1A^0\ 1A^{0}$†
1793▲	1.17	$6A^2\ 6A^4\ 1A^0\ 1A^2$
1808▲	1.23	$6A^2\ 6A^2\ 1A^0\ 1A^{0}$†
1803▲	1.37	$6A^3\ 6A^x\ 1A\ 1A^1$
1789▲	1.38	$6A^2\ 6A^2\ 1A^0\ 1A^{0}$†
1797▲	1.55	$6A^3\ 6A^3\ 1A^1\ 1A^1$
1790	1.63	$6A^2\ 6A^2\ 1A^0\ 1A^{0}$†
1805	1.66	$6A^2\ 6A^4\ 1A\ 1A^0$
1809	1.75	$6A^2\ 6A^3\ 1A^0\ 1A^{1*}$
1799▲	1.90	$6A^2\ 6A^3\ 1A^0\ 1A^{1*}$
1788▲	2.09	$6A^2\ 6A^3\ 1A^0\ 1A^{1*}$
1798	2.31	$6A\ 6A^2\ 1A\ 1A^0$
1812▲	2.41	$6A^2\ 6A^3\ 1A^0\ 1A^{1*}$
1811▲	2.52	$6A\ 6A^2\ 1A\ 1A^0$
1804▲	2.72	$6A\ 6A^2\ 1A^0\ 1A^x$
1791	3.14	$6A^2\ 6A^4\ 1A\ 1A^0$

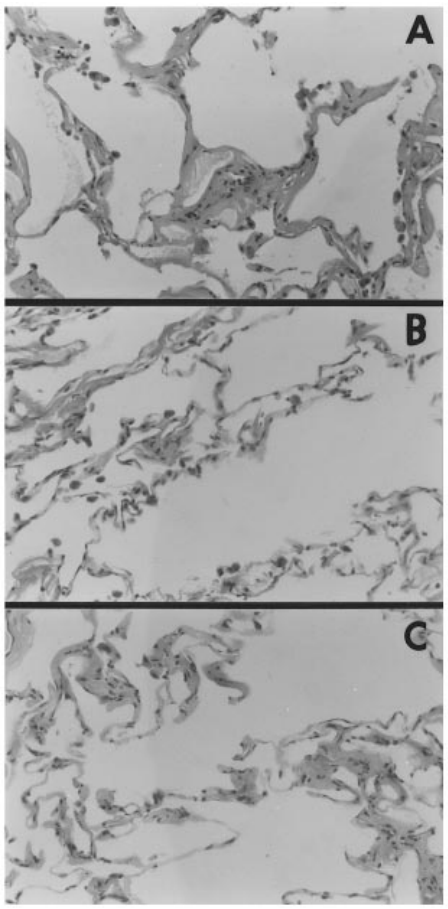

Figure 2 Histological sections of representative lung specimens

Frozen and paraffin-embedded tissue sections were cut from specimens from each group representing (**A**) low, #1801; (**B**) moderate, #1797; (**C**) high, #1804, levels of total SP-A mRNA. No major pathology was detected that would indicate the presence of chronic or acute lung disease. The histologic analysis also indicates that specimens represent comparable levels of the lung. Sections were stained with haematoxylin and eosin. Magnification × 197.

determined. Figure 1A shows the genomic structure of SP-A1 and SP-A2 and the position and orientation of oligonucleotides used in the studies described below.

The majority of the specimens used were taken from lung tissue resected to remove a localized tumor. Frozen and paraffin-embedded tissue from 15 of the specimens (marked with a triangle in Table 2) was examined histologically. The tissues did not show any major pathology and represented comparable levels of peripheral lung, based upon the size of small airways present. Figure 2 shows haematoxylin- and eosin-stained sections from three of the specimens that contained low (1801, panel A), moderate (1789, panel B) and high (1804, panel C) levels of SP-A mRNA, as determined by Northern-blot analysis (see below).

Genotype analysis

The genotypes of 23 individuals were determined by hybridization with allele-specific oligonucleotides. The sequences of these oligonucleotides were based on the SP-A alleles that have been characterized to date by both sequencing and oligonucleotide hybridizations [14]: 6A, $6A^2$, $6A^3$ and $6A^4$ for SP-A1, and 1A, $1A^0$, $1A^1$, $1A^2$ and $1A^3$ for SP-A2 [14]. The hybridization specificity of the oligonucleotides has been previously described [14]. SP-A mRNA was reverse transcribed from total lung RNA using oligonucleotide 68, common to both genes. The cDNA was then amplified using four pairs of oligonucleotides (32/293, 32/292, 314/68 and 310/68), resulting in a 5′ and a 3′ fragment for each gene (Figure 1B). The specificity of the amplification of the SP-A1 and SP-A2 fragments from reverse-transcribed RNA was verified by size difference for the 5′ fragment (Figure 1C, lanes 2–5). For the 3′ fragment, specificity was monitored by parallel amplification of characterized SP-A1 and SP-A2 cDNA clones (Figure 1C, lanes 6–11) and/or by gene-specific hybridization with oligonucleotides 351 and 433 (not shown). The amplification products were applied to slot blots and were hybridized with allele-specific oligonucleotides.

We determined the SP-A1 and SP-A2 genotypes of the 23 individuals. For one individual (1803) the SP-A1 genotype and for another individual (1804) the SP-A2 genotype could not be assigned from the hybridization pattern. This observation suggested the presence of additional alleles not yet described. All the previously characterized SP-A1 alleles and all the SP-A2 alleles, except $1A^3$ [14], were present in the study group. Ten different genotypes were identified (Table 2, column 3). The most frequent genotypes were $6A^2 6A^2 1A^0 1A^0$ (33%; $n = 7$) and $6A^2 6A^3 1A^0 1A^1$ (29%; $n = 6$). The frequencies of the SP-A1 and SP-A2 alleles in the study population were, for SP-A1: 6A, 8.9%; $6A^2$, 62.2%; $6A^3$, 22.2%; and $6A^4$, 6.7%; and, for SP-A2: 1A, 13.3%; $1A^0$, 62.2%; $1A^1$, 20.0%; and $1A^2$, 4.4%. Using the EH program for haplotype analysis [22], we detected two significant ($P < 0.01$) haplotypes, $6A^2 1A^0$ and $6A^3 1A^1$.

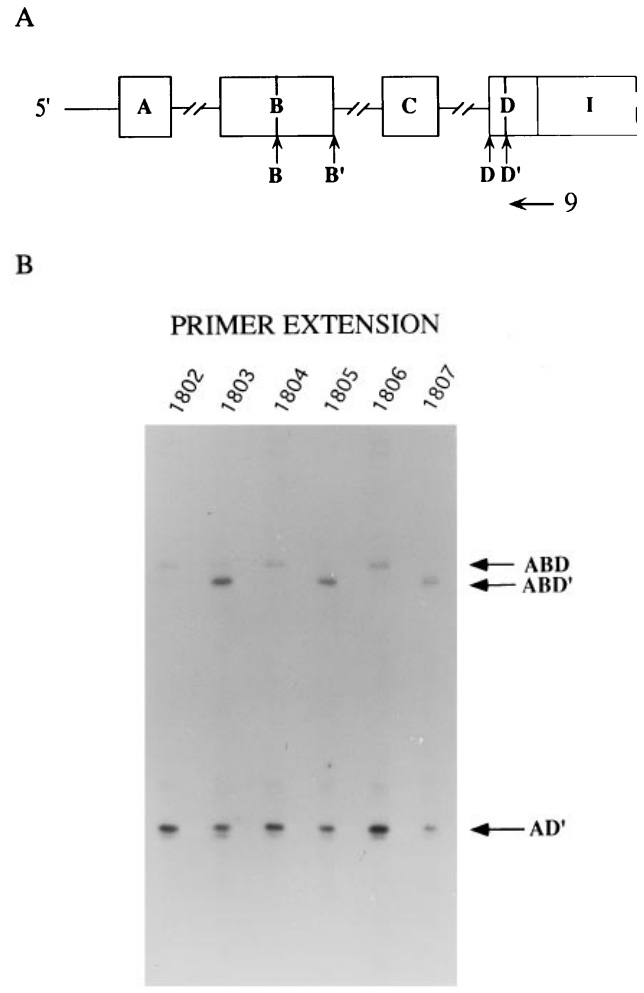

Figure 3 Primer-extension products from SP-A1 and SP-A2 mRNA

(**A**) 5′ end of the SP-A genes showing 5′ UT exons that are alternatively spliced, forming different transcripts for each SP-A gene. There are two splice acceptor sites, D and D′, three nucleotides apart at the 5′ end of exon D. Oligonucleotide 9, common to both genes, was used for primer extension. (**B**) An autoradiograph of primer-extension products from six specimens is shown. Primer extension was carried out using total lung RNA from unrelated adults, as described in the Experimental section. The major SP-A1 splice variant (AD′) and the major SP-A2 splice variants (ABD, ABD′) are marked by arrows.

Total SP-A mRNA

We used Northern blots hybridized with SP-A cDNA to compare SP-A mRNA levels among the adults studied. Duplicate blots were hybridized first with SP-A cDNA, stripped, and then hybridized with actin probe to normalize the SP-A levels. One set of filters hybridized significantly more strongly than the other. However, there was a good correlation between the relative SP-A levels on the duplicates (correlation coefficient $r = 0.83$; results not shown), so that the ranking of individuals according to total SP-A mRNA was similar on each set of filters. The relative levels of SP-A among the 23 individuals shown in Table 2 (column 2) are the means of the relative levels on the duplicate blots. The individuals are arranged in order of increasing total SP-A mRNA. (Duplicate values were not available for samples 1801, 1792 and 1808, and so these individuals were not included in statistical analyses but are included in the Table for purposes of illustration). There was a wide range of relative levels of SP-A mRNA, the difference between the lowest and the highest values being more than sixfold. Approximately 50–60% of the individuals fell in a group where there was a two-to-fourfold range in total SP-A mRNA level. The CV, a measure of the variability among the samples, was 0.49.

Individuals with the most common genotype, $6A^{2}6A^{2}1A^{0}1A^{0}$ (indicated by † in Table 2), are found mostly in the upper half of Table 2, suggesting that the level of SP-A mRNA from this genotype is relatively low (mean ± S.E. is 1.07 ± 0.13 compared with 1.70 ± 0.19 for all other genotypes). Despite the small sample number, this difference approached statistical significance ($P = 0.06$). On the other hand, the total level of SP-A mRNA in individuals with the other common genotype, $6A^{2}6A^{3}1A^{0}1A^{1}$, varies from very low to a high level, raising the possibility that this group consists of two subgroups (see the Discussion).

Relative levels of SP-A1 and SP-A2 mRNA (SP-A1/SP-A2 mRNA ratio)

We used primer extension to determine whether the relative mRNA content from each SP-A gene is the same or differs among individuals. Primer extension is a relatively direct method of measuring the relative levels of SP-A1 and SP-A2 transcripts, because the SP-A mRNA is subjected to only one manipulation, reverse transcription. Because SP-A1 and SP-A2 transcripts undergo different patterns of alternative splicing of their 5′ UT exons [11], the extension products of each gene differ in length and can be readily identified on a sequencing gel. The 5′ UT exons that undergo alternative splicing to produce different transcripts are shown in Figure 3A. The major SP-A1 transcript is AD′ and the major SP-A2 transcripts are ABD and ABD′. Figure 3B shows primer extension products of six study individuals; the SP-A1 (AD′) and SP-A2 (ABD, ABD′) products are indicated. The SP-A1/SP-A2 ratio for each individual was determined by densitometry and was calculated as [AD′/(ABD + ABD′)].

Primer extension was carried out using RNA from 21 individuals; there was insufficient RNA from individuals 1792 and 1789 for this experiment. We used radio-labelled oligonucleotide 9 (Figure 1A), common to both genes, as the primer for extension. Extension products were separated on a sequencing gel, and their sizes were verified by comparison with a sequencing ladder run in parallel lanes. The study individuals with their SP-A genotypes are listed in increasing order of SP-A1/SP-A2 ratio in Table 3. The SP-A1/SP-A2 mRNA ratio varies widely among individuals; the lowest ratio was 0.94 for individual 1794 and the highest was 6.80 for individual 1801. In general, the majority of individuals express more SP-A1 mRNA than SP-A2 mRNA (SP-A1/SP-A2 > 1). Individuals with the double homozygous $6A^{2}6A^{2}1A^{0}1A^{0}$ genotype (indicated by † in Table 3) tend to have moderate-to-high SP-A1/SP-A2 ratios, shown by their position in the lower half of the Table. Statistical analysis of the data showed that the SP-A1/SP-A2 mRNA ratio of the individuals with this genotype (mean ± S.E. $= 4.76 \pm 0.45$) was significantly higher than the mean ratio of all the other individuals (2.72 ± 0.43; $P = 0.01$). The genotype $6A6A^{2}1A1A^{0}$ (Table 3) appears to be associated with a low to moderate ratio.

Relative levels of the SP-A2 major transcripts (ABD/ABD′ ratio)

SP-A1 and SP-A2 genes undergo different patterns of alternative splicing of 5′ UT exons [11]. Earlier studies from this laboratory have reported that the relative levels of the major SP-A2 splice

Table 3 SP-A1/SP-A2 mRNA ratio and genotype

The SP-A1/SP-A2 mRNA ratio was determined for 21 individuals using primer extension as described in the Results. The individuals are listed in order of increasing SP-A1/SP-A2 mRNA ratio, with their genotypes. Individuals with the genotype $6A^2 6A^2 1A^0 1A^0$ (†) have a moderate-to-high ratio, whereas individuals with genotype 6A $6A^2$ 1A $1A^0$ (*) are associated with a moderate-to-low ratio.

Individual number	SP-A1/ SP-A2	Genotype
1793	0.94	$6A^2$ $6A^4$ $1A^0$ $1A^2$
1791	0.95	$6A^2$ $6A^4$ 1A $1A^0$
1811	1.10	6A $6A^2$ 1A $1A^{0*}$
1798	1.16	6A $6A^2$ 1A $1A^{0*}$
1803	1.41	$6A^3$ $6A^x$ 1A $1A^1$
1805	1.72	$6A^2$ $6A^4$ 1A $1A^0$
1807	1.74	6A $6A^2$ 1A $1A^{0*}$
1809	2.99	$6A^2$ $6A^3$ $1A^0$ $1A^1$
1797	3.52	$6A^3$ $6A^3$ $1A^1$ $1A^1$
1810	3.70	$6A^2$ $6A^3$ $1A^0$ $1A^1$
1800	3.87	$6A^2$ $6A^2$ $1A^0$ $1A^{0}$†
1787	4.04	$6A^2$ $6A^3$ $1A^0$ $1A^2$
1808	4.04	$6A^2$ $6A^2$ $1A^0$ $1A^{0}$†
1788	4.11	$6A^2$ $6A^3$ $1A^0$ $1A^1$
1790	4.13	$6A^2$ $6A^2$ $1A^0$ $1A^{0}$†
1804	4.28	6A $6A^2$ $1A^0$ $1A^x$
1799	4.33	$6A^2$ $6A^3$ $1A^0$ $1A^1$
1806	4.50	$6A^2$ $6A^2$ $1A^0$ $1A^{0}$†
1802	5.21	$6A^2$ $6A^2$ $1A^0$ $1A^{0}$†
1812	6.40	$6A^2$ $6A^3$ $1A^0$ $1A^1$
1801	6.80	$6A^2$ $6A^2$ $1A^0$ $1A^{0}$†

Table 4 ABD/ABD′ ratio and SP-A2 genotype

The ratio of the major SP-A2 5′ UT region splice variants, ABD/ABD′, was determined for 20 individuals using primer extension as described in the Results. The data are the means ± S.E. for the individuals of each genotype.

SP-A2 genotype	ABD/ABD′	Number of individuals
$1A^0$ $1A^2$	0.36 ± 0.18	2
1A $1A^0$	0.11 ± 0.01	5
1A $1A^1$	0.17	1
$1A^1$ $1A^1$	3.65	1
$1A^0$ $1A^1$	3.90 ± 0.30	5
$1A^0$ $1A^0$	3.89 ± 0.29	6

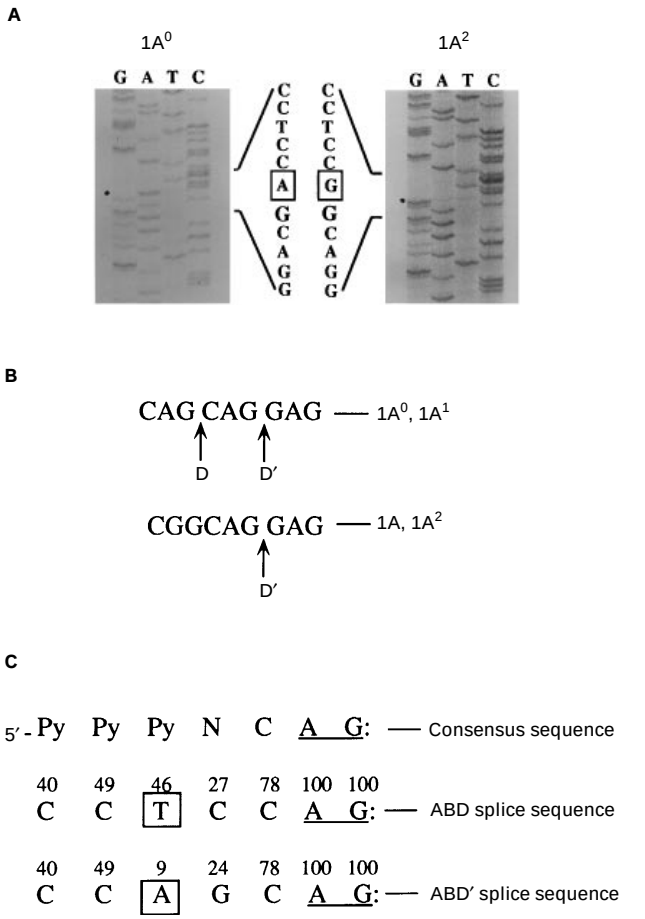

Figure 4 Sequence differences in SP-A2 alleles affect splice site selection

The genetic basis of the variation in ABD/ABD′ ratio. (**A**) The splice-recognition sequence at the 5′ end of UT exon D of two SP-A2 alleles, $1A^2$ and $1A^0$, was sequenced. The nucleotide change from A ($1A^0$) to G ($1A^2$) that eliminates the D splice site is boxed. (**B**) The intron nucleotide sequences at the 5′ end of exon D for alleles $1A^0$ and $1A^1$ and for alleles 1A and $1A^2$ are shown. Arrows mark the splice junctions, D and D′, that are used by SP-A2 alleles to give ABD (for $1A^0$ and $1A^1$ alleles) and ABD′ (for 1A, $1A^0$, $1A^1$ and $1A^2$ alleles) transcripts. (**C**) The extended intron consensus sequence of a splice junction (top) and the corresponding $1A^0$ or $1A^1$ sequence if splicing occurs at the ABD (middle) or at the ABD′ (bottom) splice site are shown. The invariant AG is underlined in each case. The numbers above the ABD and ABD′ sequences are the frequency with which that nucleotide occurs at that position in the consensus sequence. In the ABD transcripts, the nucleotide (boxed) five bases upstream from the AG splice site is T in 46% of splice junctions, and in ABD′ transcripts is A in 9% of splice junctions. Therefore the ABD pattern is favoured in $1A^0$ and $1A^1$ alleles.

variants (ABD and ABD′) differ among individuals [11,18]. The ABD and ABD′ variants arise from splicing of the A and B untranslated exons on to either the D (ABD transcripts) or D′ (ABD′ transcripts) splice-acceptor site of UT exon D. Splicing can occur at two sites on exon D in the SP-A2 gene, because of the presence of two consensus splice sequences, D and D′, at the 5′ end of exon D (Figure 3A). The variation in the ABD/ABD′ ratio was confirmed in the present study and is illustrated in Figure 3B, which shows primer extension products of both SP-A1 (AD′) and SP-A2 (ABD, ABD′) for six study individuals. The ratio of ABD to ABD′ transcripts was determined from densitometric analysis of autoradiographs of primer-extension products.

In Table 4, the study individuals are grouped according to their SP-A2 genotype. The data indicate that the ABD/ABD′ ratio varies according to the SP-A2 genotype of the individual. Individuals with SP-A2 genotype $1A^0/1A$, $1A^0/1A^2$ or $1A/1A^1$ have ratios < 1.0, i.e. there are more ABD′ than ABD transcripts, whereas individuals with genotypes $1A^0/1A^0$, $1A^0/1A^1$ and $1A^1/1A^1$ have ratios > 1.0, i.e. ABD transcripts predominate. Similar results were obtained when the samples were analysed using RT-PCR (results not shown). The correlation of genotype and ratio of SP-A2 splice variants suggests that the variability of this ratio has a genetic basis, such that 1A and $1A^2$ alleles favour splicing at the ABD′ splice site and $1A^0$ and $1A^1$ favour splicing at the ABD splice site. To determine whether sequence differences in the splice site of the various alleles could explain site selection, we sequenced the pertinent region of the 1A and $1A^2$ alleles. The published SP-A2 genomic sequence [9] is the $1A^0$ allele [11], and the intron sequence at the splice junction at the 5′ end of exon D is CAGCAG. Figure 4A shows the sequence of the splice-recognition site of the $1A^0$ and $1A^2$ alleles. As hypothesized, there was a sequence change in the splice-recognition sequence

of the $1A^2$ allele (Figure 4A). The sequence was changed from CAGCAG to CGGCAG in the $1A^2$ allele, resulting in the elimination of the D splice site, so that $1A^2$ transcripts only have the ABD′ splice pattern whereas $1A^0$ transcripts can have either the ABD or ABD′ pattern (Figure 4B). The sequence of the 1A allele was the same as $1A^2$ in this region, and the sequence of $1A^1$ was the same as $1A^0$ (results not shown).

The data in Table 4 also show that alleles $1A^0$ and $1A^1$ favour splicing at the ABD site quite strongly. This preference for the ABD site is explained after comparison of the intron sequence at the ABD and ABD′ splice sites with the corresponding extended consensus splice sequence [23]. Figure 4C makes this comparison. The numbers above the ABD and ABD′ sequences are the frequency with which that particular nucleotide occurs at that position in the consensus sequence. The invariant AG nucleotides adjacent to the splice site are underlined. The nucleotide (boxed) five bases upstream from the splice site is T (as occurs for ABD transcripts) in 46 % of splice junctions and A (as occurs in ABD′ transcripts) in 9 % of splice junctions. The preference for T in the consensus sequence explains why the ABD pattern is favoured and therefore why ABD transcripts predominate in $1A^0$ and $1A^1$ alleles, where both splice sites are available.

Table 5 SP-A1 and SP-A2 allele frequency distribution

The frequency (percent) of the SP-A1 and SP-A2 alleles in the present study population ($n = 45$ alleles for each SP-A1 and SP-A2) is compared with the frequency of a larger population ($n = 253$ alleles for SP-A1 and $n = 245$ alleles for SP-A2) [14]. The SP-A2 allele $1A^3$ was not represented in the present study population.

Alleles	Study population (%)	Larger population (%)
SP-A1		
6A	9.1	9.3
$6A^2$	63.6	54.4
$6A^3$	20.5	27.8
$6A^4$	6.8	8.5
SP-A2		
1A	13.6	11.5
$1A^0$	61.4	56.5
$1A^1$	20.5	21.0
$1A^2$	4.5	8.7
$1A^3$	0	0.8

DISCUSSION

SP-A is the major non-serum protein associated with pulmonary surfactant. A variety of different functions have been ascribed to SP-A [6], so that variation in the amount of SP-A has the potential to affect surfactant biology at a number of different sites. In the present study we have confirmed our earlier observation that the level of SP-A mRNA varies among individuals [19]. We observed a sixfold difference in total SP-A mRNA between the highest- and lowest-expressing individuals (Table 2); the high degree of variability in the group is reflected in the CV of 0.49. It is our hypothesis that the individual variability in the amount of SP-A mRNA is a reflection of one or more of the complexities observed at the level of the SP-A genes. In this study we investigated (1) variability among individuals in the level of total SP-A mRNA and in the relative levels of SP-A1 and SP-A2 mRNA as a function of SP-A genotype and (2) variability among individuals in the ratio of the SP-A2 major splice variants (ABD/ABD′) as a function of SP-A2 genotype.

Correlation of genotype with total SP-A mRNA level, SP-A1/SP-A2 mRNA ratio and the ABD/ABD′ ratio

The frequencies of the SP-A1 and SP-A2 alleles are very similar to those observed in a much larger population recently described by Floros and colleagues [14] (Table 5). This similarity suggests that the group of individuals studied here, though small, is roughly representative of the general population. Because of the small size of this study population ($n = 23$) and the large number of potential SP-A genotypes, conclusions can be drawn for only the most frequent genotypes, i.e. for $6A^26A^21A^01A^0$ and $6A^26A^31A^01A^1$.

The total SP-A mRNA level varies among individuals (Table 2). Individuals with the genotype $6A^26A^21A^01A^0$ (indicated by † in Table 2) have a low-to-moderate amount of SP-A mRNA. Because these individuals are homozygous for both SP-A genes, they must carry the $6A^21A^0$ haplotype, suggesting that total SP-A mRNA from this haplotype is low and that the amount of $6A^2$ mRNA and/or of $1A^0$ mRNA is low in these individuals. The $6A^26A^21A^01A^0$ genotype was also associated with a moderate-to-high SP-A1/SP-A2 mRNA ratio (Table 3†). Taken together, these observations indicate that the content of $1A^0$ mRNA is lower than that of $6A^2$.

The other common genotype, $6A^26A^31A^01A^1$ (indicated by * in Table 2), appears to fall into two subgroups, one with a low level of total SP-A mRNA and the other with high total SP-A mRNA, suggesting that there may be other differences that distinguish these two apparent subgroups that are not readily discernible from this study: for example, (i) one group may contain a sequence difference within the coding sequence that is not detected by the hybridization oligonucleotides used here, i.e. this group may represent another allele; (ii) there could be differences in the 3′ UT region of the allele(s) of one group that alter the stability of its mRNA, which in turn affects the level of mRNA; or (iii) there may be differences in the genetic background of the two subgroups that alter the expression of the alleles present, i.e. gene modifiers [24].

There was no correlation between total SP-A mRNA and SP-A1/SP-A2 mRNA ratio (correlation coefficient $r = 0.29$). This observation suggests that the total amount of SP-A mRNA in an individual is not determined by the relative expression of the SP-A genes. Rather, the picture that emerges is that total SP-A mRNA level and the relative levels of SP-A1 and SP-A2 mRNA produced in a given individual vary as a function of the specific alleles present, i.e. of the SP-A genotype. At the present time, the basis for the difference in expression among SP-A alleles is not known. However, the data suggest that individual alleles differ in the level of transcription or stability of their message, resulting in different amounts of mRNA.

The majority of individuals have an SP-A1/SP-A2 ratio greater than one (Table 3), indicating an excess of SP-A1 mRNA. McCormick and Mendelson [25] determined the relative levels of SP-A1 and SP-A2 mRNA in four adults. They found that SP-A2 mRNA made up 75 % of total SP-A mRNA in each case. Our results, in contrast, suggest that the relative contribution of each gene can vary considerably among individuals, but that in most cases there is more SP-A1 than SP-A2 mRNA. The reason for this apparent difference is unclear, though it is possible that it is the result of individual variation.

In addition, variability in the ratio of the major SP-A2 splice variants is also determined by genotype. Thus certain alleles (1A and $1A^2$) can only splice at the ABD′ site whereas others ($1A^0$

and 1A[1]) can use both sites, because of sequence differences within the intron splice-recognition sequence adjacent to the exon D splice junction in SP-A2 (Figure 4). The physiological significance of the splicing differences (if any) is unclear, because both ABD and ABD′ transcripts can be translated *in vitro* [11] and are also translated *in vivo* [18]. The rate of translation of each of these transcripts may differ, but, to date, no such information is available.

A wide range of SP-A levels has been measured in the bronchoalveolar lavage (BAL) of healthy control individuals in several studies, suggesting that SP-A protein level does vary among individuals [26–29]. It is not known whether the differences among individuals in BAL SP-A levels reflect an underlying range of SP-A mRNA levels, as described here. In the present study, the level of SP-A protein in the lung tissue of each individual was not measured. To date, no systematic studies of correlation between human SP-A protein and SP-A mRNA level in different individuals have been reported. However, a number of studies using human fetal lung explant culture have investigated the response of SP-A protein and its message to a variety of stimulatory and inhibitory regimens. Of particular interest was the observation of Ballard and co-workers [30] that in the explants they studied there was considerable variability in the amount of both SP-A and its message between different specimens and that the protein and message levels tended to be correlated. The observed variability in level of SP-A mRNA among specimens may have been a reflection of differences in SP-A genotype of the fetal lungs used in their experiments, as described in the present report. The results of explant studies show that, in general, the protein level tends to follow the mRNA level. For example, the content of both SP-A protein and SP-A mRNA increases with time in culture in untreated explants [30] and in explants exposed to dibutyryl cAMP, epidermal growth factor, tri-iodothyronine and interferon [31–33]. In contrast, insulin, transforming growth factor-β and retinoic acid inhibit the accumulation of both SP-A and its message in explant culture [32,34,35]. In the human adenocarcinoma cell line H441-4, phorbol ester treatment decreases the amount of both SP-A and SP-A mRNA [36]. The response to glucocorticoids of both SP-A and its message is biphasic, showing increased expression at low concentrations of glucocorticoids and inhibition of expression at high concentrations [37,38]. These results strongly suggest that the SP-A level is correlated with the level of SP-A mRNA, at least under the conditions of explant culture. It is possible that the situation in adults is different. In adults, the SP-A level may change relative to the SP-A mRNA level as a result of increased or decreased degradation of SP-A or by a change in the efficiency of the translation of SP-A mRNA into protein in different individuals. However, at this time, no information of this nature is available for SP-A in human adult lung.

The range of SP-A in healthy individuals that can maintain normal lung function has not been well defined. Nor have there been studies where the amount, or range of amounts, of functional SP-A has been determined in BAL of healthy individuals. However, the level of SP-A protein in BAL was found to be altered during certain disease states: decreased in acute respiratory distress syndrome [28] and idiopathic pulmonary fibrosis [26], and increased in sarcoidosis and hypersensitivity pneumonitis [27]. Smoking had no significant affect on the level of SP-A in BAL of healthy volunteers [27]. The differences among humans in the level of SP-A mRNA we observed in the present study, as determined by histological analysis, do not appear to be the result of lung pathology (such as chronic or acute lung disease) or to be due to different sampling sites (for example, distal versus proximal) in the lung.

Summary

Using RNA isolated from 23 unrelated adult individuals, we have shown that: (1) the level of SP-A mRNA varies among individuals; (2) one SP-A genotype associates with a low-to-moderate level of SP-A mRNA; (3) the variability in the ratio of the two major SP-A2 transcripts results from sequence differences in the UT exon D splice-recognition sequence; (4) the relative level of SP-A1 to SP-A2 mRNA (SP-A1/SP-A2 mRNA ratio) varies among individuals and the level of SP-A1 mRNA is generally greater than the level of SP-A2 mRNA; (5) one SP-A genotype associates with a moderate-to-high SP-A1/SP-A2 mRNA ratio and another with a low-to-moderate ratio; and (6) there is no correlation between the SP-A1/SP-A2 mRNA ratio and total SP-A mRNA level.

The minimum level of SP-A mRNA required to maintain normal surfactant metabolism and function is currently unknown. A level that is sufficient under normal conditions may prove inadequate in the event of disease or other stress on the respiratory system. Immunohistochemical studies of lungs of infants who died of respiratory distress syndrome showed reduced levels of surfactant proteins [39]. SP-A, in particular, was greatly reduced in or absent from these lungs [40]. A correlation between the level of SP-A and the severity of respiratory distress syndrome has also been noted [41]. Our results suggest that SP-A genotype has an impact on the level of SP-A mRNA and that one genotype is associated with relatively low amounts of SP-A mRNA. Thus, should an infant possessing a genotype that correlates with low levels of SP-A mRNA be born prematurely, it is possible that the risk that this infant will develop respiratory problems is significantly increased, compared with an infant of comparable age but with a different genotype, i.e. one that correlates with high levels of SP-A mRNA. Furthermore, it is possible that a particular stress/disease results in different outcomes in adult individuals with different genotypes.

We thank Susan DiAngelo and Gina Deiter for expert technical assistance. This work was supported by Grant HL49823 from the National Institutes of Health.

REFERENCES

1 King, R. J., Simon, D. and Horowitz, P. M. (1989) Biochim. Biophys. Acta **1001**, 294–301
2 Voss, T., Eistetter, H., Schafer, K. P. and Engel, J. (1988) J. Mol. Biol. **201**, 219–227
3 Suzuki, Y., Fujita, Y. and Kogishi, K. (1989) Am. J. Respir. Cell Mol. Biol. **140**, 75–81
4 Williams, M. C., Hawgood, S. and Hamilton, R. L. (1991) Am. J. Respir. Cell Mol. Biol. **5**, 41–50
5 Dobbs, L. G., Wright, J. R., Hawgood, S., Gonzales, R., Venstrom, K. and Nellenbogen, J. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 1010–1014
6 Weaver, T. E. and Whitsett, J. A. (1991) Biochem. J. **273**, 249–264
7 Floros, J. and Karinch, A. M. (1995) Am. J. Physiol. **268** (Lung Cell. Mol. Physiol. **12**), L162–L165
8 White, R. T., Damm, D., Miller, J., Spratt, K., Schilling, J., Hawgood, S., Benson, B. and Cordell, B. (1985) Nature (London) **317**, 361–363
9 Katyal, S. L., Singh, G. and Locker, J. (1992) Am. J. Respir. Cell Mol. Biol. **6**, 446–452
10 Voss, T., Melchers, K., Scheirle, G. and Schafer, K. P. (1991) Am. J. Respir. Cell Mol. Biol. **4**, 88–94
11 Karinch, A. M. and Floros, J. (1995) Am. J. Respir. Cell Mol. Biol. **12**, 77–88
12 McCormick, S. M., Boggaram, V. and Mendelson, C. R. (1994) Am. J. Physiol. **266** (Lung Cell. Mol. Physiol. **10**), L354–L366
13 Rishi, A., Hatzis, D., McAlmon, K. and Floros, J. (1992) Am. J. Physiol. **262** (Lung Cell. Mol. Physiol. **6**), L566–L573
14 Floros, J., DiAngelo, S., Koptides, M., Karinch, A. M., Rogan, P., Nielson, H., Spragg, R., Watterberg, K. and Dieter, G. (1996) Am. J. Respir. Cell Mol. Biol. **15**, 489–498
15 Floros, J., Steinbrink, R., Jacobs, K., Phelps, D., Kriz, R., Sultzman, L., Recny, M., Jones, S., Taeusch, H. W., Frank, H. A. and Fritsch, E. F. (1986) J. Biol. Chem. **261**, 9029–9033

16 Floros, J. and Karinch, A. M. (1995) in Surfactant Therapy for Lung Disease (Robertson, B. and Taeusch, H. W., eds.), pp. 95–106, Marcel Dekker, New York
17 Krizkova, L., Sakthivel, R., Olowe, S. A., Rogan, P. and Floros, J. (1994) Am. J. Physiol. **266** (Lung Cell. Mol. Physiol. **10**) L519–L527
18 Karinch, A. M. and Floros, J. (1995) Biochem. J. **307**, 327–330
19 Floros, J., Phelps, D. S., deMello, D. E., Longmate, J., Harding, H., Benson, B. and White, T. (1991) Exp. Lung Res. **17**, 91–104
20 Hatzis, D., Deiter, G., deMello, D. E. and Floros, J. (1994) Exp. Lung Res. **20**, 57–72
21 Floros, J., Phelps, D. S., Harding, H. P., Church, S. and Ware, J. (1989) Am. J. Physiol. **257** (Lung Cell. Mol. Physiol. **1**), L137–L143
22 Terwillinger, J. D. and Ott, J. (1994) in Handbook of Human Genetic Linkage, pp. 188–210, The Johns Hopkins University Press, Baltimore
23 Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. and Sharp, P. A. (1986) Annu. Rev. Biochem. **55**, 1119–1150
24 Rames, G. and McKusik, V. A. (1994) Nat. Genet. **7**, 451–453
25 McCormick, S. M. and Mendelson, C. R. (1994) Am. J. Physiol. **266** (Lung Cell. Mol. Physiol. **10**), L367–L374
26 McCormack, F. X., King, T. E., Voelker, D. R., Robinson, P. C. and Mason, R. J. (1991) Am. Rev. Resp. Dis. **144**, 160–166
27 Hamm, H., Luhrs, J., Rotache, J. G.y., Costabel, U., Fabel, H. and Bartsch, W. (1994) Chest **106**, 1766–1770
28 Gregory, T. J., Longmore, W. J., Moxley, M. A., Whitsett, J. A., Reed, C. R., Fowler, A. A. F., Hudson, L. D., Maunder, R. J., Crim, C. and Hyers, T. M. (1991) J. Clin. Invest. **88**, 1976–1981
29 Honda, Y., Takahashi, H., Shijubo, N., Kuroki, Y. and Akino, T. (1993) Chest **103**, 496–499
30 Ballard, P. L., Hawgood, S., Liley, H., Wellenstein, G. A., Benson, B. and White, R. T. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 9527–9531
31 Odom, M. J., Snyder, J. M. and Mendelson, C. E. (1987) Endocrinology **121**, 1155–1163
32 Whitsett, J. A., Weaver, T. E., Lieberman, M. A., Clark, J. C. and Daugherty, C. (1987) J. Biol. Chem. **262**, 7908–7913
33 Ballard, P. L., Liley, H. G., Gonzales, L. W., Odom, M. W., Ammann, A. J., Benson, B., White, R. T. and Williams, M. C. (1990) Am. J. Respir. Cell Mol. Biol. **2**, 137–143
34 Dekowski, S. A. and Snyder, J. M. (1992) Endocrinology **131**, 669–676
35 Metzler, M. D. and Snyder, J. M. (1993) Endocrinology **133**, 1990–1998
36 Pryhuber, G. S., O'Reilly, M. A., Clark, J. C., Hull, W. M., Fink, I. and Whitsett, J. A. (1990) J. Biol. Chem. **265**, 20822–20828
37 Liley, H. G., White, R. T., Benson, B. J. and Ballard, P. L. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 9096–9100
38 Boggaram, V., Smith, M. E. and Mendelson, C. R. (1989) J. Biol. Chem. **264**, 11421–11427
39 deMello, D. E., Phelps, D. S., Patel, G., Floros, J. and Lagunoff, D. (1989) Am. J. Pathol. **134**, 1285–1293
40 deMello, D. E., Heyman, S., Phelps, D. S. and Floros, J. (1993) Am. J. Pathol. **142**, 1631–1640
41 Hallman, M., Merritt, T. A., Akino, T. and Bry, K. (1991) Am. Rev. Resp. Dis. **144**, 1376–1384

Received 3 June 1996/12 August 1996; accepted 29 August 1996