Differential Regulation of Human Surfactant Protein A Genes, SFTPA1 and SFTPA2, and Their Corresponding Variants

Joanna Floros1,2* and Nikolaos Tsotakos3

1 Department of Pediatrics, The Pennsylvania State University College of Medicine, Hershey, PA, United States, 2 Department of Obstetrics and Gynecology, The Pennsylvania State University College of Medicine, Hershey, PA, United States, 3 School of Science, Engineering, and Technology, The Pennsylvania State University - Harrisburg, Middletown, PA, United States

The human SFTPA1 and SFTPA2 genes encode the surfactant protein A1 (SP-A1) and SP-A2, respectively, and they have been identified with significant genetic and epigenetic variability including sequence, deletion/insertions, and splice variants. The surfactant proteins, SP-A1 and SP-A2, and their corresponding variants play important roles in several processes of innate immunity as well in surfactant-related functions as reviewed elsewhere [1]. The levels of SP-A have been shown to differ among individuals both under baseline conditions and in response to various agents or disease states. Moreover, a number of agents have been shown to differentially regulate SFTPA1 and SFTPA2 transcripts. The focus in this review is on the differential regulation of SFTPA1 and SFTPA2 with primary focus on the role of 5′ and 3′ untranslated regions (UTRs) and flanking sequences on this differential regulation as well molecules that may mediate the differential regulation.

Keywords: transcription regulation, variants, posttranscriptional regulation, promoter, surfactant protein A (SP-A)

1 INTRODUCTION

As a way of background, pulmonary surfactant, a lipoprotein complex, is essential for life. It prevents alveolar lung collapse by lowering the surface tension at the air-liquid interface of the lung alveolus. Lung alveoli are the distal airspaces in the lung, lined by epithelial Type I and Type II cells. Under normal conditions macrophages are the only immune cells present in the alveolar space. The alveolus is covered by a thin liquid layer, called hypophase. Surfactant is found at the surface of the hypophase, i.e., at the air-liquid interface, as well as in the hypophase as a surfactant reservoir. Through its ability to maintain lung alveolar stability, surfactant enables the lung to carry out its key function of O₂/CO₂ exchange. Deficiency of surfactant in prematurely born infants and dysfunction of surfactant in adults can potentially lead to serious breathing problems including death.

Pulmonary surfactant is composed of about 90% of lipids, primarily phospholipids and four non-serum proteins, the surfactant protein A (SP-A), SP-B, SP-C, and SP-D. SP-B and SP-C are hydrophobic proteins and are involved in activities that primarily affect the surfactant function. For example, surfactant is found in the form of a monomolecular surface film at the air-liquid interface of the alveolus, which is responsible for the reduction of surface tension and thus prevention of lung collapse, and it is also found as surfactant reservoir in the hypophase. The two surfactant compartments are
interconnected. During a breath there is reorganization of surfactant layers and the hydrophobic proteins are key for surfactant multilayer connection and bringing lipids from the hypophase to the air-liquid interface (1, 2). SP-A and SP-D are hydrophilic proteins and both belong to the collectin family of proteins. These are primarily involved in innate immunity, regulation of inflammatory processes and may serve as a link to adaptive immunity. In addition, SP-A contributes to various aspects of surfactant structure such as in the formation of tubular myelin (an extracellular structural form of surfactant), and the reorganization of surfactant in the hypophase. SP-D may play a role in surfactant homeostasis (1).

SP-A was the only known surfactant protein at the early times of clinical surfactant replacement trials (3, 4). The success of a human clinical study in 1980 (5), where pulmonary surfactant derived from cow lung (a natural source) was used successfully to treat prematurely born babies at risk for respiratory problems, and the failure of previous studies where off the shelf lipids were used (6), raised interest in the study of SP-A and led to the discovery of the other surfactant proteins (7–15). Although the surfactant proteins in the early years were known by various names, a nomenclature, used today, was agreed upon soon after their discovery (16). SP-A, the focus in this review, in addition to its surfactant-related functions, plays a role in the lung innate immune response and regulation of inflammatory processes (1).

Unlike rodents that have a single gene, humans (17, 18) and primates (19) have two genes, the result of gene duplication about 26.5 million years ago (19). The human SFTPA locus has been mapped at q22-q23 of chromosome 10 (20–22). This locus consists of two functional genes, SFTPA1 and SFTPA2, in opposite transcriptional orientation with a pseudogene, SFTPA3P (23), in reverse orientation relative to SFTPA1 at about 15kb away from the 5′ region of SFTPA1 (21). The two genes are in linkage disequilibrium (24). The SFTPA locus also includes another pseudogene, the MBL3P (mannose binding lectin family member A3 pseudogene) (Figure 1A) (25). Although one study found the SFTPA locus close to the MBP locus (26), another study found that the MBP locus is located at a large distance, at about 25,000-35,000kb, from the SFTPA locus (21). Radiation hybrid mapping has placed SFTPA2 and SFTPDP (another surfactant protein gene) on the 5′ region of SFTPA1 at about 40 and 120 kb, respectively. Their orientation relative to the centromere is SFTPD-SFTPA2-SFTPA3P-SFTPA1-telomere. Although the evolutionary advantage of the SFTPA gene duplication has not been studied, one can only speculate. The SP-A protein is shown to be relatively conserved through the major vertebrate groups and it has been proposed that the surfactant system is an evolutionary prerequisite for air-breathing species (27). We postulate that the original role of SP-A was surfactant-related and at some time in evolution was “co-opted” to serve in host defense. Perhaps with the dual role, the SFTPA gene was subject to evolutionary selection that led to gene duplication. The available literature, as discussed elsewhere (1) indicates that for the most part both gene products carry similar functions but one seems to do a better job in host defense activities and the other in surfactant-related activities. Nonetheless, the functional complementation of the two protein products may mean that the gene duplication in the primate lineage was followed by subfunctionalization via selective pressure that keeps both genes functional (28). This is in contrast with another host defense gene, the mannose binding lectin (MBL), which underwent pseudogenization and lost its second functional gene in humans (29). At present, this hypothesis is simply speculative.

The SFTPA1 and SFTP2A genes encode proteins that contain both collagenous and carbohydrate regions (22, 30) that places them in the family collagenous C-type lectins or collectins. SP-D, another surfactant protein encoded by the SFTPD gene, also contains both collagenous and carbohydrate regions and it is placed in the same family of proteins (31), along with the mannose binding protein (22). Collectins are soluble pattern recognition receptors and are part of the innate immune system (32–35). They bind various molecules on the surface of microorganisms, such as carbohydrate containing structures and lipids and may eliminate microorganisms employing various mechanisms. Collectins may also modulate regulation of inflammatory or allergic processes and the adaptive immune system. In addition, they play a role in the clearance of apoptotic cells (36).

The human surfactant protein A genes have been identified with extensive genetic and epigenetic variability in coding and non-coding regions and this variability has been associated with many pulmonary diseases (37–43). The impact of the coding genetic variability on function has been reviewed elsewhere (1). Briefly, the most frequently observed coding variants in the general population for each SFTPA gene are six (1A, 1A′, 1A2, 1A3, 1A4) for SP-A2 and four (6A, 6A2, 6A3, 6A4) for SP-A1 (30, 44). However, the focus in the present review is primarily on the role of untranslated and flanking regions on the regulation of human SFTPA genes. The 5′ flanking regions of the SFTPA1 and SFTP2A genes share some conserved cis-acting regulatory elements with differing degrees of sequence conservation. Such elements have been shown to play a role in the differential regulation of the two genes, via certain transcription factors, such as NKX2.1/TTF-1 and NF-kB, or transcription coactivators such as the CBP/p300 factors. These may explain the differential regulation under different stimuli, such as in the presence of dexamethasone or CAMP analogs (45–47). In addition, epigenic regulation can partially explain the differential regulation of the two genes. DNA methylation sites have been identified upstream of the transcription start site (TSS) of both SFTP2A and SFTP2A (48, 49), while histone modifications have been implicated in the regulation under certain stimuli (50).

The human surfactant protein A is extensively co- and post translationally modified resulting in a large number of isoforms as shown by two-dimensional gel electrophoresis. Following the use of metabolic inhibitors and enzymes, this complex group of isoforms is reduced to a small number of isoforms that coincide with the isoforms of primary translation products of human lung RNA (51–53). The cloning of the genomic SFTP1A and SFTP2A sequences (17, 18) and of their cDNAs (52), enabled comparison of the two groups of sequences and this comparison revealed 5′-untranslated region (UTR) splice variability and 3′-UTR sequence variability in the SFTP1A and SFTP2A transcripts, as discussed below, as well as sequence variability within the coding region, but this has been reviewed elsewhere (1).
In humans, multiple transcripts have been identified that are due to 5′-UTR splice variability (54). Although in other species a single transcript (55) or more than one transcript (19, 27) have been identified, it is not known whether these are due to differences in splicing or polyadenylation. Karinch et al. (54), employing primer extension in an attempt to map the transcription start site of each human SFTPA gene and 5′ RACE (rapid amplification of cDNA ends) to fully characterize their transcripts, using RNA from two unrelated individuals, discovered an extensive 5′-UTR splice variability as shown in Figure 1B (upper panel). Exon C′, not shown in Figure 1, and a similar 5′-UTR splice variability was also described by McCormick et al. with a different terminology (56). In this review, we use the Karinch terminology because this has been largely used in subsequent publications. SFTPA1 transcripts appear to use A, A′ and A′′ transcription start sites of “exon A” with equal frequency but the SFTPA2 transcripts use only the A site of “exon A”. The significance of this (if any) has not been addressed further (54).

The major 5′-UTR variants for SFTPA2, ABD and ABD′, are distinguished from the major SFTPA1 5′-UTR variant by the inclusion of exon B (eB). The difference between D and D′ is 3 nucleotides with the D′ having the additional 3 nucleotides, the result of a splice site favorability between D and D′ due a single nucleotide change (57). Two minor 5′-UTR splice variants were observed for SFTPA1 (ACD′ and AB′D) plus some rare (not shown in Figure 1B) variants (54). The transcripts from each gene carrying a different major or minor 5′-UTR splice variant are translated both in vitro (except for the AB′D′) (54) and in vivo as shown by polysome bound RNA (58). However, differences in both relative translatability and relative levels of the splice variants were observed among individuals (58).

Extensive sequence variability including small deletions/insertions was also observed in 3′-UTR (54, 56, 59, 60). This
variability included an 11-nucleotide (11-nt) insertion/deletion (Figure 1C) that was initially described for an SP-A1 variant named 6A1 (59). This 6A1 variant was identical to the most frequently found SFTPA1 variant, 6A2, except at a single nucleotide (54). However, since the 6A1 variation was not present in subsequent sequencing data, it was presumed that this was a sequencing error and the 6A1 is referred to as 6A2 thereafter. This 11nt sequence provides potential binding sites for miRNA (61), shown to regulate expression (62). In addition to the 11-nt, other elements have been identified, by sequence comparison, in the 3′-UTR of SFTPA1 and SFTPA2. These include the minimum AU-rich element motif UUAUUUAUU shown elsewhere to mediate mRNA degradation (63). This is present in SFTPA2 variants at position 926-935 but not in the SFTPA1 variants (61).

2 REGULATION OF SFTPA

Under baseline conditions, SFTPA mRNA levels vary significantly among individuals (64) with a sixfold difference between high and low expression among individuals (57). The lack of correlation between total SFTP mRNA levels and the SFTPA1/SFTPA2 transcript ratio indicated that the levels in an individual may vary as a function of the SFTPA genotype, where the level of transcription and/or stability of mRNA may differ among variants (57). A variability in SP-A protein levels in bronchoalveolar lavage is observed among individuals (65–68) and during development (69). Together these indicate that there may be mechanisms that differentially affect the expression of the SP-A variants. In fact, in response to a variety of stimulatory or inhibitory regimens in fetal lung explants or in the human adenocarcinoma H441 cell line, the levels of human SP-A protein and/or mRNA change significantly (70–73). Also, SP-A levels may change in certain disease states (65–67, 74–76). Furthermore, inhibitory or stimulatory substances were shown to differentially affect the regulation of SFTPA1 and SFTPA2 mRNA in fetal lung explants or cell lines (46, 77–79). Although differences in protein and mRNA levels in health and in various disease states have been observed, to the best of our knowledge no comprehensive study has been done to correlate mRNA levels of each human SFTPA gene/variant and protein levels in samples from patients or healthy individuals or experimental systems where each transcript in its entirety is studied. We speculate that given the opportunity for complex regulation, as discussed in this review, it is probably unrealistic to think that a direct correlation of the overall mRNA and protein levels could exist without providing additional specific information. Such information may include reference to the specific genetic variant, the splice variant, the specific conditions, i.e., exposure to various insults including environmental stressors, and other. Experimental models of 5′-UTR or 3′-UTR regions of the SFTPA1/A2 variants are shown to exhibit differences in response to various insults and thus the mRNA/protein levels of the two genes and/or of their variants may differ. It would probably require additional reagents and approaches, such as gene- and variant-specific antibodies, and direct RNA sequencing to make such a correlation meaningful.

3 THE ROLE OF 5′ AND 3′ UNTRANSLATED REGIONS AND FLANKING SEQUENCES IN THE DIFFERENTIAL REGULATION OF SFTPA1 AND SFTPA2 TRANSCRIPTS

3.1 Impact of 5′-UTR Splice Variants in the Differential Regulation of SFTPA1 and SFTPA2

These variants have been shown to regulate several steps/processes in SP-A regulation (80). They differentially affect translation and mRNA stability as assessed by in vitro transient expression of reporter gene constructs containing different 5′-UTR (A′D′, ABD, AB′D′ and A′CD′) splice variants. All variants compared to control vectors had a positive effect on gene expression as shown by increases in reporter gene activity and mRNA levels, with the ABD performing significantly better than the rest. In terms of the translation efficiency index (reporter activity/mRNA) a differential effect was observed by the splice variants. Compared to the control, both ABD and AB′D′ exhibited higher translation efficiency whereas the other two splice variants, A′D′ and A′CD′, exhibited a lower efficiency. Algorithms predicting the secondary structure stability of the 5′-UTRs revealed that, compared to others (A′D′, AB′D′, A′CD′), the ABD structure was the most energetically favored one. Furthermore, the ABD was shown to exhibit a lower rate of mRNA decay upon inhibition of transcription with actinomycin D. Collectively, these indicate that the ABD splice variant has a better secondary structure stability and a lower rate of mRNA decay (80).

Splice variants (ABD, A′D′, A′B′D′ with the exception of A′ CD′) were shown to differentially mediate internal ribosome entry site (IRES) activity i.e., cap-independent translation with the ABD exhibiting the highest IRES activity and A′D′ the next highest whereas the AB′D′ and A′CD′ exhibited low or no IRES (81). Secondary structure stability and especially the presence of a double loop structure in ABD and A′D′ (but absent in AB′D′ and A′CD′) as well as cis acting elements (in ABD) and perhaps other factors may all differentially contribute to the cap-independent translation. The ABD IRES activity was responsive to specific environmental stressors (i.e., to diesel PM but not to ozone exposure). Furthermore, the double-loop structure, which is important in cap-independent translation, didn’t seem to be necessary for cap-dependent translation activity, as shown with the A′D′ splice variant (81).

One major difference between the major SFTPA1 and SFTPA2 5′-UTR splice variants is the presence or absence of exon B (eB). The eB presence in the UTR, as discussed above, results in a better outcome, whether mRNA stability, rate of mRNA decay, secondary structure and perhaps other, indicating that eB may be an important regulatory element. In a series of reporter gene constructs or in vitro translation experiments, eB was found to be an enhancer of transcription, if placed upstream of heterologous 5′-UTR or in its natural 5′-UTR, as it increased mRNA content regardless of position or orientation (82). eB also increased translation of mRNA reporter transcripts in the
presence or absence of poly-A, when placed within its natural sequence environment but in heterologous 5′-UTR increased translation only in the presence of poly-A (82).

The 14-3-3 proteins form homo- or heterodimers and by binding a variety of ligands including kinases, phosphatases, transmembrane receptors, etc., regulate a variety of functions, including cell cycle control, translation, apoptotic cell death, other (83–85). eB interacts either alone or within the context of the surrounding 5′-UTR sequences with 14-3-3 proteins. RNA pulldown assays, RNA affinity chromatography and surface plasmon resonance analyses showed that eB binds directly most of the 14-3-3 protein isoforms (β, γ, ε, η, σ, τ/θ) except isoform zeta (ζ). The latter isoform may bind eB indirectly because isoform zeta was identified by mass spectroscopy of shift and pull-down assays to be part of the eB-protein complex. Regardless of its presence in the eB-protein complex isoform zeta does not affect SP-A2 levels upon inhibition by shRNA knocked down (86). However, inhibition of the other 14-3-3 eB binding isoforms resulted (except isoform ñ) in a downregulation of SP-A2 without any change in SP-A1 levels. Isoform ñ did not show any gene-specific downregulation, as the levels of both SP-A1 and SP-A2 were negatively affected. Furthermore, differences in the stability of eB/14-3-3 isoform complexes have been observed (86). Deletion and mutation mapping analyses revealed two regulatory motifs in eB, GUCCGUGAU (next to exon A) and GGAGCCUGAA (near exon D) that are important for protein binding as assessed by shift assays (87). The eB RNA/protein complexes, one major and one minor contain in addition to the 14-3-3 proteins a number of other proteins that include, among others, ribosomal, cytoskeletal and translation factor proteins (87). Competition experiments with excess AD or ABD RNA of the eB-mediated shifts did not disrupt the eB shifts entirely (as the eB RNA competitor did) but resulted in altered mobility shifts with a lower size. The collective observations of the eB shifts competed with AD or ABD excess RNAs along with the mass spectroscopy data of the identity of the proteins in the eB-shifts before and after competition are summarized in a schematic representation of Figure 2 [adapted from (87)]. The 14-3-3 proteins surprisingly were not competed with the ABD RNA but these were indeed competed with the AD RNA competitor. The reasons for this are not clear. The 14-3-3 proteins were present in shifts with either the eB or ABD probe but not with the AD. It was postulated that since in silico analysis showed that 6nt at the 3′ end of exon A were part of an eB regulatory element, one possibility is that the ABD but not the AD provided some kind of stability at the junction of A-B resulting in a partial displacement.

The ACD 5′ UTR splice variant has been described as a minor splice variant of SP-A1 transcripts and is found only in SP-A1 transcripts (54, 56). The exon C of this splice variant is 60 nucleotides long and contains two upstream AUG (uAUG) sites in addition to the primary (p) AUG (Figure 3). The AUG closer to the TATA box is in frame with the pAUG whereas the other one is not. The in-frame uAUG results in an N-terminally...
Tsotakos and colleagues (88) showed that the uAUGs in the nucleotides downstream. Using a variety of approaches, the pAUG, is followed with an in-frame stop codon, eight ACD

Extended isoform, but the additional residues do not seem to alter the site of cleavage by the signal peptidase (88). The out-of-frame uAUG introduces an ORF that overlaps with the primary ORF (the stop codon is within the coding region of SFTPA1, corresponding to residues 72-73 of the protein product of the main ORF). The production of any peptide from the overlapping ORF and its effects (cis or trans) on SP-A1 protein production have yet to be evaluated. An uAUG has also been described for exon B (Figure 3). However, this uAUG, although in frame with the pAUG, is followed with an in-frame stop codon, eight nucleotides downstream. Using a variety of approaches, Tsotakos and colleagues (88) showed that the uAUGs in the ACD’ splice variant decrease SP-A1 expression without affecting the size of the mature protein. The ACD’ transcripts appear to be present in the majority of individuals and their expression can be affected by mechanical injury. Their contribution to the SP-A1 transcript pool may be regulated by different stimuli including LPS and dexamethasone. Moreover, the SP-A1 AD’ (major) 5′ splice variant and the SP-A1 ACD’ (minor) 5′-UTR variant may be differentially regulated (88). Interestingly, the presence of exons C or B’ in the ACD’ and AB’/D’ splice variants of the SFTPA1 transcript, respectively, may introduce G-quadruplex structures in the 5′-UTR, which may affect translation initiation (89, 90). Such structures are absent from all SFTPA2 splice variants, as analyzed by the online QGRS Mapper tool (91).

There are five polymorphisms in the 5′-UTR of SFTPA1 that are categorized as possibly loss of function (pLoF) by the Genome Aggregation Database (gnomAD, v.3.1.1) (92), out of a total of 271. Specifically, rs1317624468 and rs1020324172 may affect splicing of the 5′-UTR exons, but the frequency of either variant is very low. This means that there is low confidence in these variants. Similarly, in the 5′-UTR of SFTPA2, there are 12 pLoF out of a total of 290 polymorphisms. As is the case with variants in SFTPA1, the identified short nucleotide variants (SNVs) seem to be associated with splice donor/acceptor sites, but the confidence on the effect these variants have is low, given their low frequencies. The SNVs were identified with the use of the UCSC Genome Browser (93).

### 3.2 3′-UTR-Mediated Regulation of Human SFTPAs

A number of SFTPA1 and SFTPA2 transcripts have been identified. These encode different protein variants i.e., with differences in their coding region, and these protein variants are discussed elsewhere (1). The 3′-UTRs of the transcripts of these protein variants, in transient transfection experiments compared to control vector, have been shown to differentially reduce mRNA and protein levels, as assessed by the activity of the reporter gene at baseline and in response to dexamethasone treatment (94). The inhibition in response to dexamethasone is glucocorticoid specific as both dexamethasone and hydrocortisone decreased reporter gene activity (95). Dihydrocortisone and phorbol ester 12-O-tetradecanoylphorbol-13-acetate on the other hand did not have any effect on reporter gene activity. The former was shown previously not to regulate SP-A (47) and the latter to affect SP-A regulation at the transcriptional level (96).

The 3′-UTRs of the SFTPA1 transcripts encoding protein variants, 6A³, 6A³, 6A⁴, exhibit a differential effect on translation with no significant difference found between the 3′-UTRs of the two studied SFTPA2 protein variants, 1A⁰ and 1A³ (61). An 11-nt element, described in the introduction (59), is located at position 405. This element that is present in the 3′-UTR of all the SFTPA2 3′-UTR sequences investigated to date and in the SFTPA1 transcript encoding the 6A³ protein variant but absent in other SFTPA1 transcripts studied, had a negative impact on translation. Upon its removal, translation increased, and the stability of the predicted secondary structure was changed. In silico analysis of the 11-nt element revealed seven potential miRNA binding sites (61). miRNAs are small noncoding RNAs that regulate gene expression at the posttranscriptional level via interactions with untranslated mRNA sequences.

Three miRNAs (miRNA-183, miRNA-4495 and miRNA-612) with potential binding sites within or near the 11-nt sequence (Figure 1C), via the use of miRNA mimics and/or antagonirs, were shown to inhibit gene expression of all 3′-UTR-constructs that included the 11-nt element (i.e., all SFTPA2 transcripts and the SFTPA1 transcript of the 6A³ protein

---

**Human SP-A 5′ UTR translation start sites**

![Diagram of Human SP-A 5′ UTR translation start sites](image-url)

**FIGURE 3** | Translation start sites at the 5′-UTR. The primary translation start site (pAUG) is marked with a black arrow. Other upstream translation start sites (uAUG) in frame with the pAUG are marked with a red arrow. X denotes a stop codon in frame with the immediately uAUG.
variant (62). One miRNA (miRNA-4507) negatively affected the reporter gene activity of SFTPA1 transcripts that lacked the 11-nt sequence, and another (miRNA-767) inhibited expression of both SFTPA1 and SFTPA2 transcripts. Collectively, these data indicate that miRNA regulatory pathways are involved in the SP-A regulation. This has been further validated with the knockdown of Drosha, an important effector of miRNA maturation. Inhibition of Drosha in primary human alveolar type II cells via siRNA, resulted in an increase in the levels of SP-A (97).

In summary, both 5′-UTR and 3′-UTR are important in the regulation of the human SFTPAs. Transient transfection experiments of reporter genes showed that 5′-UTR and 3′-UTR have an additive effect on translation. In addition, the poly-A tail also contributes to SFTP regulation (61). Transcripts of constructs containing SFTPA2 5′-UTR variants in the presence or absence of poly-A, displayed a higher level of in vitro translation products than SFTPA1 5′-UTR (AD′). Moreover, the presence of the poly-A tail, even in the absence of 3′-UTR, increased translation (61).

### 3.3 SP-A Flanking Sequences in the Regulation of SP-A1 and SP-A2

The 5′-flanking regions of the SFTPA genes have been studied in multiple species. Sequence comparisons between rat and human SFTPA genes identified one proximal (up to 225bp upstream of TSS) and one distal (-1115bp in rats/-938bp in humans) conserved element in the 5′-flanking regions (98). This conservation led to further exploration of the flanking regions for regulatory elements. Sequencing analysis of these regions in rat, rabbit, baboon and human (19, 21, 98, 99) led, as described below, to the identification of several cis elements. These elements act as binding sites for transcription factors during both basal expression and in response to cellular signals.

#### 3.3.1 Promoter Analysis in Animal Species

Developmental studies using rabbit and baboon as model identified a DNase I-hypersensitivity site -180 to -80 bp of the TSS after gestational day 21 in rabbits and day 140-160 in baboons, indicating potential changes in the proximal promoter region around the developmental timing of gene activation (100). Primer extension analysis of the upstream sequence in rabbits revealed an octamer that is one nucleotide different from the consensus cAMP response element (CRE) at -261bp (100). Interestingly, this CRESP-A fails to bind the CREB transcription factor or a basic leucine zipper polypeptide, indicating that a different transcription factor may be responsible for binding to this specific site (101). Similar studies followed the characterization of the two SFTPA genes in baboons (19). Regulation patterns in the presence of dibutyryl-cAMP and dexamethasone in the baboon were similar to the ones observed in the rabbit model (102) prompting further study of the flanking sequences.

Analysis of the rat sequences and CAT reporter assays identified a silencing element between base pairs -195 and -163, which was bound by members of the C/EBP family of transcription factors (103, 104). Further analysis of the rabbit SP-A promoter by fusion of different promoter sites with the human growth hormone (hGH) structural gene, used as reporter, revealed potential binding sites for several transcription factors, such as Sp1 at -190bp and AP-1 at -416 and -255bp. Four elements homologous to glucocorticoid response element (GRE) were also identified (99). In addition, two E-box sequences, one proximal and one distal, were identified. These seem to be bound by homo- and heterodimers of the Upstream Stimulatory Factors 1 and 2 (USF1 and USF2), playing a role in basal and hormonal regulation of the rabbit SFTP gene (105, 106). An E-box motif was also identified in positions -8 to -3 of the murine SP-A gene promoter, but it was considered to not play a lung-specific regulatory role as assessed by transfection of MLE-15 cells, a cell line derived from lung tumors produced in transgenic mice expressing SV40 large T antigen driven by the lung-specific human SP-C promoter (107). Furthermore, DNase I fingerprinting assays and EMSA with bacterially expressed TTF-1/Nkx2.1 revealed three binding sites that comprise a TTF1-binding element (TBE) for each baboon gene (108). A similar element with four binding sites was discovered in the murine SFTP promoter (107). In the rat SFTP gene promoter, there is an insertion in positions -316 to -211 that is considered to have occurred after the divergence of the mouse and rat lineages. Within the conserved sequences, there are five potential TTF-1 binding sites, of which at least four were present in protected regions by DNase I fingerprinting analysis (55, 109). The glucocorticoid inhibition of SP-A expression was found to be mediated by TTF-1 (110). The activity of TTF-1 is dependent on phosphorylation by Protein Kinase A (PKA), a cAMP-induced kinase (111). During development, TTF-1 expression depends on the presence of certain microRNAs (112) and the Hepatocyte Nuclear Factor 3B (HNF3-β). The latter belongs to the winged family of transcription factors and targets other genes critical for the differentiation of respiratory epithelial cells (113). PKA activated by cAMP increases TTF-1 phosphorylation and binding to the TBE (111) but also enhances the interaction of TTF-1 with the CREB-interacting protein (CBP) and the steroid receptor coactivator 1 (SRC-1, a different member of the nuclear receptor coactivator family) (114). The TBE also contains a reverse-oriented NF-κB binding site (115). Interleukin-1 treatment, along with TTF-1, increases NF-κB binding to the TBE (115). Dexamethasone increases expression of the NF-κB inhibitor, IκB-α, thus blocking its transcriptional activity (110), while increasing recruitment of histone deacetylases 1 and 2 near the TBE (Figure 4A, highlighted by the blue ribbon), as shown by chromatin immunoprecipitation (116). NF-κB seems to be the common mediator of the developmental timing of expression of genes that are involved in lung innate immunity and surfactant homeostasis, as indicated by time-dependent transcriptome profiling in two strains of mice (117). Two more factors co-regulate SP-A expression along with TTF-1, at least in mice; GATA-6, which binds to a GATA-binding site at positions -69 to -64 of the murine gene promoter (118) and B-Myb, which binds to an element in positions -380 to -371 (119). High throughput ChIP-seq analysis of TTF-1/Nkx2.1 confirmed its binding to...
genes critical for lung function and health (120, 121). Although a TBE containing three TTF-1 binding sites was also identified in the human SFTPA2 as a cAMP-responsive cis element in studies using transgenic mice, only one of the binding sites (the middle one) is identical between the baboon and human SFTPA2 (122).

3.3.2 SFTPA1 and SFTPA2

Hormonal regulation studies by transfection of lung cell lines of human origin with WT and CRE-mutated constructs verified the responsiveness of the CRE to the cAMP analogue dibutyryl-cAMP (Bt2cAMP) (99). Dexamethasone inhibits the cAMP-induced expression of the reporter gene despite stimulating transcription of SFTPA in the absence of cAMP in the same system, indicating an interaction between the CRE and a glucocorticoid response element. Critically, the two human genes are differentially regulated by a number of signaling molecules, such as cAMP, glucocorticoids, and insulin (46, 77, 79), a fact that highlights the importance of the promoters as elements that mediate differential expression of the two genes. Such studies are further complicated by the fact that the dexamethasone effect in SP-A gene expression is biphasic and dose-dependent (47, 70, 123).

Subsequent studies focused mostly on regulatory elements of hSFTPA2, as it was found to be more responsive to stimulatory effects by cAMP than hSFTPA1 (79). Studies using transfected type II cells identified a CRE element in the hSFTPA2 gene (124), although similar studies have not been performed for the SFTPA1 gene. The CRE element is bound by the Estrogen-Related Receptors, ERα and ERγ, but only the ERα receptor increased hSFTPA transcription, while the ERγ had no effect (125). A mechanistic study confirmed that the action of ERRα is mediated by PKA and revealed SRC-2 (steroid receptor coactivator 2) as a cofactor of the cAMP-induced transcriptional activation of SP-A (126). SRC-2 is downregulated by the glucocorticoid receptor (127), thus partially explaining the dexamethasone-induced mitigation of cAMP-dependent SFTPA2 transcription. Moreover, the CRE (or ERRE, Estrogen-Related Receptor Element, Figure 4A) works cooperatively with a GT box (128) to mediate basal and cAMP-induced changes in SP-A expression. Based on supershift experiments, the GT box is bound by at least five protein complexes, two of which contain Sp1, a ubiquitously expressed transcription factor (128).

With regards to inhibitory stimuli, hSFTPA1 was found to be more responsive to dexamethasone treatment than hSFTPA2 in H441 cells (78, 79) and in fetal lung explants treated with 100nM dexamethasone, the expression of SFTPA1 mRNA was inhibited to a greater degree than SFTPA2 (77). Study of the mechanistic aspects of this inhibition identified the -32/+63 region (relative to the TSS) as the dexamethasone response element, indicated by a yellow band in Figure 4A (45). Removal of the region -227/-31, encompassing the CRE/ERRE, GT box, and TBE (111, 124, 128), did not affect the dexamethasone response, but it significantly attenuated the basal transcriptional hSFTPA1 promoter activity, indicating that these elements are not involved with this dexamethasone-mediated regulatory pathway (45). Another inhibitory agent for both human genes is phorbol ester (96). Deletion analysis to identify promoter elements that are responsible for this effect identified a region downstream of the TTS (+309/+329 in hSFTP1). A member of the Jun, but not of the Fos family of proteins, was identified by supershift assays as a binding transcription factor to this site (129). Furthermore, the +318/+324 region contains a sequence (TGACTGA)
similar to the AP-1 consensus binding site (shown as “AP-1 binding site” in Figure 4A), thus implicating AP-1 complexes in the transcription of the SFTPA genes.

A functional study employing CRISPRi, a method using a fusion of deactivated Cas9 and the repressor KRAB, thereby inhibiting genes without deleting them, was performed on identified targets of TTF-1/Nkx2.1, including SFTPA1 and SFTPA2 (130). Prior ChIP-seq analysis had indicated that TTF-1 binds to a proximal upstream region of SFTPA1 and a distal upstream region of both SFTPA1 and SFTPA2 (131, 132). Targeting either of these sites with CRISPRi in A549 and H441 cells suppressed the expression of SFTPA1 significantly, but SFTPA2 was suppressed only by the sgRNA targeting the distal region (130). Deletion of the distal region, which is about 20kbp away from either gene’s TSS and is located close to the SFTPA3P pseudogene, by the CRISPR/Cas9 approach in A549 repressed the expression of both SFTPA1 and SFTPA2 (130), which may indicate a regulatory role of the pseudogene.

The presence of two surfactant protein genes with high degree of similarity in the flanking regions makes it technically challenging to discern mechanisms of differential regulation. The combination of models, such as cell lines and animal models has led however to the identification of several transcription factor binding sites (Figure 4). At the very least, ERRα, USP1/2 heterodimers, Sp1, TTF-1/Nkx2.1 and AP-1 complexes have been shown to regulate SP-A transcription, but any differential effects have yet to be shown unequivocally. Advances in gene editing techniques may elucidate this topic in the future.

4 EPIGENETIC REGULATION

Alterations of SFTPA in lung cancer have been observed in various experimental settings (133–139). In a few cases, SFTPA was shown to be useful as a marker of differential diagnosis of metastatic lung cancer and mesotheliomas (138, 140–143). A high-throughput approach identified aberrant methylation of CpG sites of several genes to associate with lung cancer (144) indicating that this may be a contributing process in lung cancer. A subsequent DNA CpG methylation profiling of the SFTPA1 gene promoter identified two CpG SFTPA1 sites (SFTPA1_370 and SFTPA1_1080, Figure 4C) to be hypomethylated in lung cancer (adenocarcinoma and squamous cell carcinoma). In normal lung tissue the level of methylation of another SFTPA1_1468 CpG site (not shown to significantly differ in its methylation content in cancer lung tissue) was associated with the level of SFTPA1 transcripts. This CpG is located 160 nucleotides upstream of the TATAA box. The high level of unmethylated CpG_1468 was correlated with a high level of SFTPA1 transcripts indicating that the methylation status of this site may play a role in SP-A1 expression. This site is absent from SFTPA2. Of relevance, rare SFTPA1 transcripts, including a more frequently found SFTPA1 transcript coding for the 6A2 protein variant, were shown to associate with risk for lung cancer (145). A CpG DNA site methylation difference was also observed between normal and cancer lung tissue for the SFTPA2 gene promoter (49). This CpG site is located -2215 upstream of the transcription start site (Figure 4B) and exhibited a higher level of methylation in lung cancer especially in adenocarcinoma. Moreover, the level of SFTPA2 mRNA and protein were reduced in lung cancer whereas the mRNA level of DNA methyltransferases (DNMT1 and DNMT2) was increased. An in-silico analysis revealed a number of potential binding sites of transcription factors around this CpG methylation site indicating that apart from its potential as a marker, this CpG site modification may interfere with the binding of regulatory factors affecting SFTPA2 expression.

DNA methylation, which is affected by environmental factors, air pollution, smoking, diet among others (146–150), may be a regulatory mechanism for SFTPA gene expression. The differential regulation of the methylation status of specific CpGs in SFTPA1 and SFTPA2 maybe one of the epigenetic processes that does not only apply to lung cancer but also to other health states (151). Furthermore, epigenetic phenomena that modulate changes in gene function without changing the nucleotide sequence include several processes, such as DNA methylation, histone modifications, miRNAs, and splice variants. The role of miRNAs, splice variants and DNA methylation has been discussed above. Histone acetylation and methylation have been shown to affect SFTPA expression in the lung during development and hypoxia (50, 116, 152). The developmental timing of SFTPA1 expression is associated with enhanced acetylation and decreased methylation of histone H3 at the SFTPA promoter. Histone methyltransferases Suv39H1 and Suv39H2 are bound to the TBE prior to induction of SFTPA by cAMP (152). Their transcript levels are inversely correlated with the developmental pattern of SP-A expression (152). Increased O2 tension facilitates the induction of histone H3 acetylation on lysines 9 and 14 at the TBE, while hypoxia induces dimethylation of lysine 9 of H3 (50) and recruitment of Suv39H1 and Suv39H3 to the TBE (152). Dexamethasone treatment increased nuclear levels of the histone deacetylases HDAC-1 and HDAC-2, but not the total levels of histone H3, in human fetal type II cells (116). ChIP analysis indicated that dexamethasone also increases the occupancy of the TBE specifically by HDAC-1 and HDAC-2. The cAMP analogue Bt2cAMP increased the levels of acetylated and phosphorylated histone H3; this effect was antagonized by dexamethasone, which promoted demethylation of H3K9 globally and locally, in the TBE region of the SFTPA promoter (116).

Although differential allele expression has not been studied for the human SFTPA genes either in the lung or extrapulmonary tissues, it is worth noting that this mechanism may be operative under certain conditions, as it is shown for the rat SFTPA under unperturbed conditions (153). SFTPA was shown under these conditions to exhibit a balanced biallelic expression in the lung but in colon was both balanced and imbalanced and family studies indicated that inheritable factor(s) may contribute to the regulation of differential allele expression (153).

5 CONCLUSION/DIscussion

Understanding differences between human SFTPA1 and SFTPA2 genes and their corresponding variants is both useful and important as such information could be implemented in personalized regimens. For example, in a pilot study where precision cut lung slices from human donor lungs were treated with varying
concentrations of methylprednisolone, a pharmacologic relationship was observed between treatment and SFTPA genotype (154). As lung donors and recipients are treated with immunosuppressive regimens, one may in future pharmacogenetic studies of lung transplants further investigate as well as consider immunosuppressive regimes tailored according to the donor’s genetic background. Recent developments in sequencing technologies, particularly the advent of long-read sequencing, could contribute to our understanding of variant frequencies and their potential regulation. Furthermore, identification of key regulators, either cis or trans, can be useful in modulating specific SFTPA1 or SFTPA2 gene expression as it may be appropriate i.e., in the case of the prematurely born infants where levels of SP-A are low and/or in the course of infection. Of interest, the SFTPA2 transcript coding for the 1A protein variant has been shown to exhibit a protective effect in terms of survival in both animal models after infection and/or other insults (155, 156) and in lung transplant patients (39). Thus, in such cases it could be advantageous to modulate expression of SP-A1 and/or SP-A2 proteins.

6 OUTSTANDING ISSUES AND FUTURE STUDIES

The work presented in this review indicates that the human SFTPA genes are under extensive and complex regulatory control that merits further experimentation. Our current knowledge on this topic is based for the most part on studies where different regions of a given gene was studied i.e., flanking region, 5′-UTR, 3′-UTR. So at the present time it is difficult to know exactly how a given variant will respond to a stimulus or under a certain environmental condition in its entirety. For example, regulatory mechanisms of one region (i.e., 5′-UTR) may attenuate or enhance i.e., an up or down regulation imparted by mechanisms operative in another region (i.e., 3′-UTR, flanking region) or even nullify effects. A major advance to our understanding could be achieved if the entire SFTPA locus is studied as a unit in order to better assess the commonalities and differences between the two genes as well as the potential role of the SFTPA3P on the regulation of the functional genes, as pseudogenes may contribute to the regulation of their functional counterpart (157–159). It is currently unknown whether SFTPA3P is expressed or not, and if so, under what conditions and at what developmental stage. If it is expressed, it could act as a lincRNA, thus regulating the neighboring parental genes in a number of possible ways (160). Given that the two functional genes are in opposite transcriptional orientation and may share regulatory cis elements that may or may not work in a coordinated regulation, examining the entire locus as a whole would be a sensible approach, albeit one with substantial challenges. Identifying an appropriate study system would be a start. Given the size of the locus, generation of humanized mice would be technically challenging, so studies in non-human primates might be more appropriate.

AUTHOR CONTRIBUTIONS

JF had the idea for the article. JF and NT performed the literature search and wrote the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Floros J, Thorenoor N, Tsotakos N, Phelps DS. Human Surfactant Protein SP-A and SP-A2 Variants Differentially Affect the Alveolar Microenvironment, Surfactant Structure, Regulation and Function of the Alveolar Macrophage, and Animal and Human Survival Under Various Conditions. Front Immunol (2021) 12:2889. doi: 10.3389/fimmu.2021.681639

2. Floros J, Phelps DS. Pulmonary Surfactant. In: TL Yaksh, editor. Anesthesia: Biologic Foundations. Philadelphia, PA: Lippincott-Raven (1997). p. 1259–79.

3. King RJ, Klass DJ, Gikas EG, Clements JA. Isolation of Apoproteins From Canine Surface Active Material. Am J Physiol (1973) 224(4):788–95. doi: 10.1152/ajplegacy.1973.224.4.788

4. King RJ, Ruch J, Gikas EG, Platzker AC, Creasy RK. Appearance of Paoproteins of Pulmonary Surfactant in Human Amniotic Fluid. J Appl Physiol (1975) 39(5):735–41. doi: 10.1152/jappl.1975.39.5.735

5. Fujisawa T, Maeta H, Chida S, Morita T, Watabe Y, Abe T. Artificial Surfactant Therapy in Hylaine-Membrane Disease. Lancet (1980) 1 (8199):55–9. doi: 10.1016/S0140-6736(80)90489-4

6. Noutsios GT, Floros J. Highlights of Early Pulmonary Surfactant: Research From Bench to Clinic. Pneumon (2013) 26(4):350–4.

7. Phelps DS, Smith LM, Taeusch HW. Characterization and Partial Amino Acid Sequence of a Low Molecular Weight Surfactant Protein. Am Rev Respir Dis (1987) 135(5):1112–7. doi: 10.1164/arrd.1987.135.5.1112

8. Phelps DS, Harding HP. Immunohistochemical Localization of a Low Molecular Weight Surfactant-Associated Protein in Human Lung. J Histochem Cytochem (1987) 35(11):1339–42. doi: 10.1177/35.11.3309049

9. Jacobs KA, Phelps DS, Steinbrink R, Fisch J, Kriz R, Mitsock L, et al. Isolation of a Cdnna Clone Encoding a High Molecular Weight Precursor to a 6-Kda Pulmonary Surfactant-Associated Protein. J Biol Chem (1987) 262(20):9808–11. doi: 10.1016/S0021-9258(18)48005-2

10. Hawgood S, Benson BJ, Schilling J, Damm D, Clements JA, White RT. Nucleotide and Amino Acid Sequences of Pulmonary Surfactant Protein SP 18 and Evidence for Cooperation Between SP 18 and SP 28-36 in Surfactant Lipid Adsorption. Proc Natl Acad Sci USA (1987) 84(1):66–70. doi: 10.1073/pnas.84.1.66

11. Glasser SW, Kerfghan TR, Weaver T, Pilot-Matias T, Fox JL, Whitsett JA. Cdnna and Deduced Amino Acid Sequence of Human Pulmonary Surfactant-Associated Proteolipid SPL(Phe). Proc Natl Acad Sci USA (1987) 84(12):4007–11. doi: 10.1073/pnas.84.12.4007

12. Fisher JH, Shannon JM, Hofmann T, Mason RJ. Nucleotide and Deduced Amino Acid Sequence of the Hydrophobic Surfactant Protein SP-C From Rat: Expression in Alveolar Type II Cells and Homology With SP-C From Other Species. Biochim Biophys Acta (1989) 995(3):225–30. doi: 10.1016/0005-2760(89)90040-X

13. Whitsett JA, Hull WM, Ohning B, Ross G, Weaver TE. Immunologic Identification of a Pulmonary Surfactant-Associated Protein of Molecular Weight = 6000 Daltons. Pediatr Res (1986) 20(8):744–9. doi: 10.1203/00006450-198605000-00009

14. Whitsett JA, Ohning BL, Ross G, Meuth J, Weaver T, Holm BA, et al. Hydrophobic Surfactant-Associated Protein in Whole Lung Surfactant and its Importance for Biophysical Activity in Lung Surfactant Extracts Used for Replacement Therapy. Pediatr Res (1986) 20(5):460–7. doi: 10.1203/00006450-198605000-00016

15. Persson A, Chang D, Rust K, Moxley M, Longmore W, Crouch E. Purification and Biochemical Characterization of CP4 (SP-D), a Collagenous Surfactant-Associated Protein. Biochemistry (1989) 28(15):6361–7. doi: 10.1021/bi00441a031
16. Possmayer F. A Proposed Nomenclature for Pulmonary Surfactant-Associated Proteins. Am Rev Respir Dis (1988) 138(4):990–8. doi: 10.1164/ arron.138.4.990

17. White RT, Damm D, Miller J, Speart K, Schilling J, Hawgood S, et al. Isolation and Characterization of the Human Pulmonary Surfactant Apoprotein Gene. Nature (1983) 317(6035):361–3. doi: 10.1038/317361a0

18. Khatr SL, Singh G, Lockyer J. Characterization of a Second Human Pulmonary Surfactant-Associated Protein SP-a Gene. Am J Respir Cell Mol Biol (1992) 6(4):466–52. doi: 10.1165/ajrcmb.6.4.446

19. Gao E, Wang Y, McCormick SM, Li J, Seidner SR, Mendelson CR. Characterization of Two Baboon Surfactant Protein a Genes. Am J Physiol (1996) 271(4 Pt 1):L617–30. doi: 10.1152/ajplung.1996.271.AL617

20. Bruns G, Stroh H, Veldman GM, Latt SA, Floros J. The 35 Kd Pulmonary Surfactant-Associated Protein is Encoded on Chromosome 10. Hum Genet (1987) 76(1):58–62. doi: 10.1007/BF00283051

21. Hoover RR, Floros J. Organization of the Human SP-a and SP-D Loci at 10q22.2-23.1. J Biol Chem (1991) 266(2):131–8. doi: 10.1016/j.jbc.1991.02.059

22. Sastry K, Herman GA, Day L, Deignan E, Bruns G, Morton CC, et al. The Human Mannose-Binding Protein Gene: Exon Structure Reveals Its Evolutionary Relationship to a Human Pulmonary Surfactant Gene and Localization to Chromosome 10. J Exp Med (1989) 170(4):1175–89. doi: 10.1084/jem.170.4.1175

23. Korfhagen TR, Glasser SW, Bruno MD, McMahan MJ, Whitsett JA. A Portion of the Human Surfactant Protein a (SP-a) Gene Locus Consists of a Pseudogene. Am J Respir Cell Mol Biol (1991) 4(5):463–9. doi: 10.1165/ajrcmb.4.5.463

24. Liu W, Bentley CM, Floros J. Study of Human SP-a, SP-B and SP-D Loci: Allele Frequencies, Linkage Disequilibrium and Heterozygosity in Different Races and Ethnic Groups. BMC Genet (2003) 4:13. doi: 10.1186/1471-2156-4-13

25. Guo N, Mogues T, Weremowicz S, Morton CC, Sastry K. The Human Ortholog of Rhesus Mannose-Binding Protein-a Gene Is an Expressed Pseudogene That Localizes to Chromosome 10. Mamm Genome (1998) 9(3):246–9. doi: 10.1002/(SICI)1537-2001(1998)9:3<246::AID-MGEO9>3.0.CO;2-2

26. Kolbke K, Lu J, Mole SE, Kaluz S, Reid KB. Assignment of the Human Pulmonary Surfactant Protein D Gene (SFTPD) to 10q22–Q23 Close to the Surfactant Protein a Gene Cluster. Genomics (1993) 17(2):294–8. doi: 10.1006/geno.1993.1324

27. Sullivan LC, Daniels CB, Phillips ID, Orgeig S, Whitsett JA. Conservation of Surfactant Protein a: Evidence for a Single Origin for Vertebrate Pulmonary Surfactant. J Mol Evol (1998) 46(2):131–8. doi: 10.1007/PL00006287

28. Iizugi S, Liberales DA. Subfunctionalization of Duplicated Genes as a Transition State to Neofunctionalization. BMC Evol Biol (2005) 5:258. doi: 10.1186/1471-2148-5-28

29. Seyfarth J, Garred P, Madsen HO. The ‘Involution’ of Mannose-Binding Lectin. Hum Mol Genet (2005) 14(19):2859–69. doi: 10.1038/hmg/ddi318

30. Floros J, Hoover RR. Genetics of the Hydrophilic Surfactant Proteins a and D. Biochim Biophys Acta (1998) 1408(2-3):312–22. doi: 10.1016/S0005-2104(98)00077-5

31. Crouch E, Rust K, Veile R, Donis-Keller H, Grosso L. Genomic Organization of Human Surfactant Protein D (SP-D): SP-D Is Encoded on Chromosome 10q22.2-23.1. J Biol Chem (1993) 268(4):2976–83. doi: 10.1074/jbc.268.4.2976

32. Holmskov U, Thiel S, Jensens JC. Collections and Ficolins: Humoral Lectins of the Innate Immune Defense. Annu Rev Immunol (2003) 21:547–78. doi: 10.1146/annurev.immunol.21.1.40954

33. Crouch E, Wright JR. Surfactant Proteins a and D and Pulmonary Host Defense. Annu Rev Physiol (2001) 63:521–54. doi: 10.1146/annurev.physiol.63.1.521

34. Gupta G, Suroit A. Collectins: Sentinels of Innate Immunity. Bioessays (2007) 29(5):452–64. doi: 10.1002/bies.20573

35. van de Wetering JK, van Golde LM, Batenburg JJ. Collectins: Players of the Innate Immune System. Eur J Biochem (2004) 271(7):1229–49. doi: 10.1111/j.1432-1033.2004.04040.x

36. Vandiver RW, Ogden CA, Fadok VA, Hofmann PR, Brown KK, Botto M, et al. Role of Surfactant Proteins a and D in the Clearance of Apoptotic Cells In Vivo and In Vitro: Calreticulin and CD91 as a Common Collector Receptor Complex. J Immunol (2002) 169(7):3978–86. doi: 10.4049/jimmunol.169.7.3978

37. DiAngelo S, Lin Z, Wang G, Phillips S, Ramet M, Luo J, et al. Novel, Non-Radioactive, Simple and Multiplex PCR-Crpl Methods for Genotyping Human SP-a and SP-D Marker Alleles. Dis Markers (1999) 15(4):269–81. doi: 10.1153/1999/961430

38. Floros J, Kala P. Surfactant Proteins: Molecular Genetics of Neonatal Pulmonary Diseases. Am Rev Respir Dis (1998) 60:365–84. doi: 10.1146/annurev.physiol.60.1.365

39. Silveira P, Floros J. Genetic Complexity of the Human Surfactant-Associated Proteins–SP-A1 and SP-A2 Gene. J Biol Chem (2012) 531(2):126–32. doi: 10.1016/j.jbc.2012.09.111

40. DiAngelo S, Lin Z, Wang G, Phillips S, Ramet M, Luo J, et al. Novel, Non-Radioactive, Simple and Multiplex PCR-Crpl Methods for Genotyping Human SP-a and SP-D Marker Alleles. Dis Markers (1999) 15(4):269–81. doi: 10.1153/1999/961430

41. Hoover RR, Thomas KH, Floros J. Glucocorticoid Inhibition of Human SP-A1 Promoter Activity in NCI-H441 Cells. Biochem J (1999) 340: (Pt 1):71–6. doi: 10.1042/bj99460069

42. McCormick SM, Mendelson CR. Human SP-A1 and SP-A2 Genes Are Differentially Regulated During Development and by Camp and Glucocorticoids. Am J Physiol (1994) 266(1 Pt 1):L367–74. doi: 10.1152/ajplung.1994.266.1.L367

43. Liley HG, White RT, Benson BJ, Ballard PL. Glucocorticoids Both Stimulate and Inhibit Production of Pulmonary Surfactant Protein a in Fetal Human Lung. Proc Natl Acad Sci USA (1988) 85(23):9096–100. doi: 10.1073/pnas.85.23.9096

44. Lin Z, Thomas NJ, Bibikova M, Seifart C, Wang Y, Guo X, et al. DNA Methylation Markers in Pulmonary Surfactant: Evidence for a Role in Pulmonary Surfactant-Associated Protein. J Biol Chem (2001) 276(12):10031–6. doi: 10.1074/jbc.006070200

45. Lacaze-Masmonteil T, Fraslon C, Bourbon J, Raymondjean M, Kahn A. Characterization of the Rat Pulmonary Surfactant Protein a Promoter. Eur J Biochem (1999) 262(3):613–23. doi: 10.1011/j.1432-1033.1999.tb16966.x
56. McCormick SM, Boggaram V, Mendelson CR. Characterization of Mrna Transcripts and Organization of Human SP-A1 and SP-A2 Genes. Am J Physiol (1994) 266(4 Pt 1):L354–66. doi: 10.1152/ajplung.1994.266.4.L354

57. Karinch AM, deMello DE, Floros J. Effect of Genotype on the Levels of Surfactant Protein a Mrna and on the SP-A2 Splice Variants in Adult Humans. Biochem J (1997) 321(1):39–47. doi: 10.1042/bj3210039

58. Karinch AM, Floros J. Translation In Vivo of 5' Untranslated Region Splice Variants of Human Surfactant Protein-a. Biochem J (1995) 307(Pt 2):327–30. doi: 10.1012/bj3070327

59. Rishi A, Hatzis D, McAlmon K, Floros J. An Allelic Variant of the 6A Gene for Human Surfactant Protein a. Am J Physiol (1992) 262(5 Pt 1):L566–73. doi: 10.1152/ajplung.1992.262.5.L566

60. Krikova L, Sakthivel R, Olowe SA, Rogan PK, Floros J, Human SP-A: Genotype and Single-Strand Conformation Polymorphism Analysis. Am J Physiol (1994) 266(5 Pt 1):L519–27. doi: 10.1152/ajplung.1994.266.5.L519

61. Silveyra P, Wang G, Floros J. Human SP-A1 (SFTPA1) Variant-Specific 3' Utrs and Poly(a) Tail Differentially Affect the In Vitro Translation of a Reporter Gene. Am J Physiol Lung Cell Mol Physiol (2010) 299(4):L1523–34. doi: 10.1152/ajplung.00113.2010

62. Silveyra P, DiAngelo SL, Floros J. An 11-Nt Sequence Polymorphism at the 5'UTR of Human SFTPA1 and SFTPA2 Gene Variants Differentially Affect Gene Expression Levels and Mrna Regulation in Cell Culture. Am J Physiol Lung Cell Mol Physiol (2014) 307(1):L116–19. doi: 10.1152/ajplung.00739.2013

63. Zubiaga AM, Belasco JG, Greenberg ME. The Nonamer UUAUUUUAAU Is the Key AU-Rich Sequence Motif That Mediates mrna Degradation. Mol Cell Biol (1995) 15(4):2219–30. doi: 10.1128/MCB.15.4.2219

64. Floros J, Phelps DS, deMello DE, Longmate J, Harding H, Benson B, et al. The Utility of Postmortem Lung for DNA Studies; Variability and Correlation of the Expression of Surfactant Proteins in Human Lung. Exp Lung Res (1991) 17(1):91–104. doi: 10.3109/01902491909633284

65. McCormack FX, King TE Jr., Voelker DR, Robinson PC, Mason RJ. Idiopathic Pulmonary Fibrosis. Abnormalities in the Bronchoalveolar Lavage Content of Surfactant Protein a. Am Rev Respir Dis (1991) 144(1):160–6. doi: 10.1143/acr/144.1.160

66. Hamm H, Luhrs J, Guzman y Rotaech J, Costabel U, Bartsch W. Elevated Surfactant Protein a in Bronchoalveolar Lavage Fluids From Sarcoidosis and Hypersensitivity Pneumonitis Patients. Chest (1994) 106(6):1766–70. doi: 10.1378/chest.106.6.1766

67. Gregory TJ, Longmore WJ, Mosley MA, Whitsett JA, Reed CR, Fowler AA 3rd, et al. Surfactant Chemical Composition and Biophysical Activity in Acute Respiratory Distress Syndrome. J Clin Invest (1991) 88(6):1796–81. doi: 10.1172/JCI115525

68. McDonald AW, Takada H, Shijsbo N, Kuroki Y, Akino T. Surfactant Protein-a Concentration in Bronchoalveolar Lavage Fluids of Patients With Pulmonary Alveolar Proteinosis. Chest (1993) 103(2):496–9. doi: 10.1378/ chest.103.2.496

69. Cau F, Pisu E, Gerosa C, Senes G, Ronchi F, Botta C, et al. Interindividual Variability in the Expression of Surfactant Protein a and B in the Human Lung During Development. Eur J Histochem (2016) 60(3):2678. doi: 10.4084/ejh.2016.2678

70. Ballard PL, Hawgood S, Liley H, Wellenstein G, Gonzales LW, Benson B, et al. Regulation of Pulmonary Surfactant Approtein SP-28-36 Gene in Fetal Human Lung. Proc Natl Acad Sci USA (1986) 83(24):9527–31. doi: 10.1073/pnas.83.24.9527

71. Odom MJ, Snyder JM. Differential Effects of Epidermal Growth Factor and Transforming Growth Factor-Beta on Synthesis of Mrna = 35,000 Surfactant-Associated Protein in Fetal Lung. J Biol Chem (1987) 262(16):7908–13. doi: 10.1016/0021-9258(87)90755-5

72. Ballard PL, Liley HG, Gonzales LW, Odom MW, Ammann AJ, Benson B, et al. Interferon-Gamma and Synthesis of Surfactant Components by Cultured Human Fetal Lung. Am J Respir Cell Mol Biol (1990) 2(2):137–40. doi: 10.1165/arcbib/2.2.137
105. Gao E, Wang Y, Alcorn JL, Mendelson CR. The Basic Helix-Loop-Helix 
Seidner SR, Smith ME, Mendelson CR. Developmental and Hormonal 
100. Chen Q, Boggaram V, Mendelson CR. Rabbit Lung Surfactant Protein a 
Frontiers in Immunology | www.frontiersin.org November 2021 | Volume 12 | Article 76671913 
Floros and Tsotakos Regulation of SP-A Genes 
110. Alcorn JL, Islam KN, Young PP, Mendelson CR. Glucocorticoid Inhibition of 
108. Li J, Gao E, Seidner SR, Mendelson CR. Differential Regulation of Baboon SP-A1 
Mol Endocrinol 
503. doi: 10.3109/01902149309064359 
66. doi: 10.3109/01902148.2014.9297577 
97. Silveyra P, Chroneos ZC, DiAngelo SL, Thomas NJ, Noutsios GT, Tsotakos 
112. Benlhabib H, Guo W, Pierce BM, Mendelson CR. The Mir-200 Family and 
122. Liu D, Yi M, Smith M, Mendelson CR. TTF-1 Response Element is Critical 
125. Liu D, Hinshelwood MM, Giguere V, Mendelson CR. Estrogen Related 
127. Sun Y, Tao YG, Kagan BL, He Y, Simons SSJr. Modulation of Transcription 
Integrated Cistromic and Expression Analysis of Ampli 
Genes Dev 
36. doi: 10.1152/ajplung.00219.2002 
322x427 (1996) 10(2):159 
70. doi: 10.1210/jbc.2008-0282 
83. doi: 10.1210/molc.2013.02.018 
27. doi: 10.1101/gad.203208.112 
111. Li J, Gao E, Mendelson CR. Cyclic AMP-Responsive Expression of the Surfactant Protein-a Gene Is Mediated by Increased DNA Binding and Transcriptional Activity of Thyroid Transcription Factor-1. J Biol Chem (1998) 273(8):4592–600. doi: 10.1074/jbc.83.25.6151–6. doi: 10.1074/jbc.272.37.23398 
110. Alcorn JL, Islam KN, Young PP, Mendelson CR. Glucocorticoid Inhibition of 
109. Alcorn JL, Gao E, Chen Q, Smith ME, Gerard RD, Mendelson CR. Genomic 
128. Young PP, Mendelson CR. A GT Box Element Is Essential for Basal and 
Biphasic Glucocorticoid Regulation of Human SP-A2 Gene in Alveolar Type II Cells: Evidence for the Binding of Lung 
347x474 PloS One (2013) 8(7):e37046. doi: 10.1164/rccm-conference.2012.185.1_MeetingAbstracts.A1216 
123. Iannuzzi DM, Ertsey R, Ballard PL. Biphasic Glucocorticoid Regulation of 
119. Bruno MD, Whitsett JA, Ross GF, Korfhagen TR. Transcriptional Regulation 
122. Liu D, Yi M, Smith M, Mendelson CR. TTF-1 Response Element is Critical 
125. Liu D, Hinshelwood MM, Giguere V, Mendelson CR. Estrogen Related 
40. doi: 10.1074/jbc.274.39.27523 
115. Islam KN, Mendelson CR. Potential Role of Nuclear Factor Kappab and 
127. Sun Y, Tao YG, Kagan BL, He Y, Simons SSJr. Modulation of Transcription 
119. Bruno MD, Whitsett JA, Ross GF, Korfhagen TR. Transcriptional Regulation 
57. doi: 10.1093/nar/gkaa1070 
24. doi: 10.1152/ajplung.1999.276.6.L917 
22. doi: 10.1152/ajplung.00375.2002 
26. doi: 10.1152/ajplung.1992.262.6.L662 
5. doi: 10.1074/jbc.M109793200 
210. doi: 10.1101/gad.203208.112 
128. Young PP, Mendelson CR. A GT Box Element Is Essential for Basal and 
123. Iannuzzi DM, Ertsey R, Ballard PL. Biphasic Glucocorticoid Regulation of 
111. Li J, Gao E, Mendelson CR. Cyclic AMP-Responsive Expression of the Surfactant Protein-a Gene Is Mediated by Increased DNA Binding and Transcriptional Activity of Thyroid Transcription Factor-1. J Biol Chem (1998) 273(8):4592–600. doi: 10.1074/jbc.83.25.6151–6. doi: 10.1074/jbc.272.37.23398 
111. Li J, Gao E, Mendelson CR. Cyclic AMP-Responsive Expression of the Surfactant Protein-a Gene Is Mediated by Increased DNA Binding and Transcriptional Activity of Thyroid Transcription Factor-1. J Biol Chem (1998) 273(8):4592–600. doi: 10.1074/jbc.83.25.6151–6. doi: 10.1074/jbc.272.37.23398 
110. Alcorn JL, Islam KN, Young PP, Mendelson CR. Glucocorticoid Inhibition of 
109. Alcorn JL, Gao E, Chen Q, Smith ME, Gerard RD, Mendelson CR. Genomic 
112. Benlhabib H, Guo W, Pierce BM, Mendelson CR. The Mir-200 Family and 
122. Liu D, Yi M, Smith M, Mendelson CR. TTF-1 Response Element is Critical 
125. Liu D, Hinshelwood MM, Giguere V, Mendelson CR. Estrogen Related 
26. doi: 10.1152/ajplung.1992.262.6.L662 
9. doi: 10.1152/ajplung.1999.276.6.L197 
110. Alcorn JL, Islam KN, Young PP, Mendelson CR. Glucocorticoid Inhibition of 
108. Li J, Gao E, Seidner SR, Mendelson CR. Differential Regulation of Baboon SP-A1 
Mol Endocrinol 
24. doi: 10.1152/ajplung.1999.276.6.L197 
110. Alcorn JL, Islam KN, Young PP, Mendelson CR. Glucocorticoid Inhibition of 
109. Alcorn JL, Gao E, Chen Q, Smith ME, Gerard RD, Mendelson CR. Genomic 
112. Benlhabib H, Guo W, Pierce BM, Mendelson CR. The Mir-200 Family and 
122. Liu D, Yi M, Smith M, Mendelson CR. TTF-1 Response Element is Critical 
125. Liu D, Hinshelwood MM, Giguere V, Mendelson CR. Estrogen Related 
127. Sun Y, Tao YG, Kagan BL, He Y, Simons SSJr. Modulation of Transcription 
119. Bruno MD, Whitsett JA, Ross GF, Korfhagen TR. Transcriptional Regulation 
5. doi: 10.1074/jbc.M109793200 
210. doi: 10.1101/gad.203208.112 
128. Young PP, Mendelson CR. A GT Box Element Is Essential for Basal and 
Biphasic Glucocorticoid Regulation of Human SP-A2 Gene in Alveolar Type II Cells: Evidence for the Binding of Lung 
347x474 PloS One (2013) 8(7):e37046. doi: 10.1164/rccm-conference.2012.185.1_MeetingAbstracts.A1216 
123. Iannuzzi DM, Ertsey R, Ballard PL. Biphasic Glucocorticoid Regulation of 
5. doi: 10.1074/jbc.M109793200 
210. doi: 10.1101/gad.203208.112 
128. Young PP, Mendelson CR. A GT Box Element Is Essential for Basal and 
Biphasic Glucocorticoid Regulation of Human SP-A2 Gene in Alveolar Type II Cells: Evidence for the Binding of Lung 
347x474 PloS One (2013) 8(7):e37046. doi: 10.1164/rccm-conference.2012.185.1_MeetingAbstracts.A1216 
123. Iannuzzi DM, Ertsey R, Ballard PL. Biphasic Glucocorticoid Regulation of 
5. doi: 10.1074/jbc.M109793200 
210. doi: 10.1101/gad.203208.112 
128. Young PP, Mendelson CR. A GT Box Element Is Essential for Basal and 
Biphasic Glucocorticoid Regulation of Human SP-A2 Gene in Alveolar Type II Cells: Evidence for the Binding of Lung
Nuclear Factors Distinct From Sp1. *Mol Endocrinol* (1997) 11(8):1082–93. doi: 10.1210/mend.11.8.9950

129. Boecker RR, Pavlovic J, Floros J. Induction of AP-1 Binding to Intron 1 of SP-A and SP-A2 Is Implicated in the Phorbol Ester Inhibition of Human Sp-a Promoter Activity. *Exp Lung Res* (2000) 26(5):303–17. doi: 10.1080/01902140048277

130. Stuart WD, Fink-Baldauf IM, Tomoshige K, Guo M, Maeda Y. Crispr-Cas9, Hoover RR, Pavlovic J, Floros J. Regulation of SP-A Genes

131. Li S, Hursting SD, Davis BJ, McLachlan JA, Barrett JC. Environmental Exposure, DNA Methylation, and Gene Regulation: Lessons From Diethylstilbestrol-Induced Cancers. *Ann N Y Acad Sci* (2003) 983:161–9. doi: 10.1111/j.1749-6632.2003.tb05971.x

132. Yang IV, Schwartz DA. Epigenetic Control of Gene Expression in the Lung. *Am J Respir Crit Care Med* (2011) 183(10):1295–301. doi: 10.1164/rccm.201010-1579PP

133. Gao X, Jia M, Zhang Y, Breitling LP, Brenner H. DNA Methylation Changes of Whole Blood Cells in Response to Active Smoking Exposure in Adults: A Systematic Review of DNA Methylation Studies. *Clin Epigenet* (2015) 7:113. doi: 10.1186/s13148-015-0148-3

134. Maugeri A. The Effects of Dietary Interventions on DNA Methylation: Implications for Obesity Management. *Int J Mol Sci* (2020) 21(22):8670. doi: 10.3390/ijms21228670

135. Silveyra P, Floros J. Air Pollution and Epigenetics: Effects on Sp-a and Innate Host Defence in the Lung. *Swiss Med Wkly* (2012) 142:w13579. doi: 10.4444/smw.2012.13579

136. Benhabib H, Mendelson CR. Epigenetic Regulation of Surfactant Protein a Gene (Sp-a) Expression in Fetal Lung Reveals a Critical Role for Suv39h Methytransferases During Development and Hypoxia. *Mol Cell Biol* (2011) 31(10):1949–58. doi: 10.1128/MCB.01063-10

137. Lin Z, Wang Y, Zhu K. Floros J. Differential Allele Expression of Host Defensin, Pulmonary Surfactant Protein-a and Osteopontin in Rat. *Mol Immunol* (2004) 41(12):1155–65. doi: 10.1016/j.molimm.2004.06.006

138. Aramini B, Geraghty P, Lederer DJ, Costa J, DiAngelo SL, Floros J, et al. Surfactant Protein a and D Polymorphisms and Methylprednisolone Pharmacogenetics in Donor Lungs. *J Thorac Cardiovasc Surg* (2019) 157(5):2109–17. doi: 10.1016/j.jtcvs.2018.12.098

139. Thorenoor N, Umstead TM, Zhang X, Phelps DS, Floros J. Survival of Surfactant Protein-A1 and SP-A2 Transgenic Mice After Klebsiella Pneumoniae Infection, Exhibits Sex-, Gene-, and Variant Specific Differences; Treatment With Surfactant Protein Improves Survival. *Front Immunol* (2018) 9:2404. doi: 10.3389/fimmu.2018.02404

140. Thorenoor N, David SP, Kala P, Ravi R, Floros Phelps A, Umstead TM, et al. Impact of Surfactant Protein-a Variants on Survival in Aged Mice in Response to Klebsiella Pneumoniae Infection and Ozone: Serendipity in Action. *Microorganisms* (2020) 8(9):1276. doi: 10.3390/microorganisms8091276

141. Hirotsune S, Yoshida N, Chen A, Barrett L, Sugiyama F, Takahashi S, et al. An Expressed Pseudogene Regulates the Messenger-RNA Stability of Its Homologous Coding Gene. *Nature* (2003) 423(6935):91–6. doi: 10.1038/nature01535

142. Lee JT. Molecular Biology: Complexity of Gene and Pseudogene. *Nature* (2003) 423(6935):26–8. doi: 10.1038/42326a

143. Ou H, Liu Q, Lin J, He W, Zhang W, Ma J, et al. Pseudogene Annexin A2 Pseudogene 1 Contributes to Hepatocellular Carcinoma Progression by Modulating Its Parental Gene ANXA2 via Mirna-376a-3p. *Dig Dis Sci* (2021) 66(11):3903–15. doi: 10.1007/s10620-020-06734-0

144. Engebritz IM, Haines JE, Perez EM, Munson G, Chen J, Kane M, et al. Local Regulation of Gene Expression by lncRNA Promoters, Transcription and Splicing. *Nature* (2016) 539(7629):8670. doi: 10.1038/nature20149

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.