

RESEARCH ARTICLE

# Identification of Novel Short C-Terminal Transcripts of Human *SERPINA1* Gene

Nerea Matamala<sup>1</sup>, Nupur Aggarwal<sup>2</sup>, Paolo Iadarola<sup>3</sup>, Marco Fumagalli<sup>3</sup>, Gema Gomez-Mariano<sup>1</sup>, Beatriz Lara<sup>4</sup>, Maria Teresa Martinez<sup>5</sup>, Isabel Cuesta<sup>6</sup>, Jan Stolk<sup>7</sup>, Sabina Janciauskiene<sup>2</sup>, Beatriz Martinez-Delgado<sup>1</sup>\*<sup>✉</sup>

**1** Molecular Genetics Unit, Instituto de Investigación de Enfermedades Raras (IIER), Instituto de Salud Carlos III (ISCIII), Madrid, Spain, **2** Department of Respiratory Medicine, Hannover Medical School, Hannover, Biomedical Research in Endstage and Obstructive Lung Disease Hannover (BREATH), Hannover, Germany, **3** Department of Biology and Biotechnologies, Biochemistry Unit, University of Pavia, Pavia, Italy, **4** Respiratory Medicine Department, Coventry and Warwickshire University Hospital, Coventry, Warwickshire, United Kingdom, **5** Pneumology Department, Hospital 12 de Octubre, Madrid, Spain, **6** Bioinformatics Unit, Instituto de Salud Carlos III (ISCIII), Madrid, Spain, **7** Department of Pulmonology, Leiden University Medical Center, Leiden, Netherlands

✉ These authors contributed equally to this work.

\* [bmartinezd@isciii.es](mailto:bmartinezd@isciii.es)



**OPEN ACCESS**

**Citation:** Matamala N, Aggarwal N, Iadarola P, Fumagalli M, Gomez-Mariano G, Lara B, et al. (2017) Identification of Novel Short C-Terminal Transcripts of Human *SERPINA1* Gene. PLoS ONE 12(1): e0170533. doi:10.1371/journal.pone.0170533

**Editor:** Pavel Strnad, Medizinische Fakultät der RWTH Aachen, GERMANY

**Received:** October 28, 2016

**Accepted:** January 5, 2017

**Published:** January 20, 2017

**Copyright:** © 2017 Matamala et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The sequence data of the described transcripts are available at NCBI GenBank with the following accession numbers: KU535895 (ST1C4), KU535894 (ST1C5), KU755451 (ST1A4) and KU755450 (ST1A5).

**Funding:** This work has been partially funded by the Instituto de Salud Carlos III ([www.isciii.es](http://www.isciii.es)) grant PI14CIII/00070 (BMD) and SEPAR (Sociedad Española de Neumología y Cirugía Torácica, [www.separ.es](http://www.separ.es)) grant 92/2014. The funders had no role in study design, data collection and analysis,

## Abstract

Human *SERPINA1* gene is located on chromosome 14q31-32.3 and is organized into three (IA, IB, and IC) non-coding and four (II, III, IV, V) coding exons. This gene produces  $\alpha$ 1-antitrypsin (A1AT), a prototypical member of the serpin superfamily of proteins. We demonstrate that human peripheral blood leukocytes express not only a product corresponding to the transcript coding for the full-length A1AT protein but also two short transcripts (ST1C4 and ST1C5) of A1AT. In silico sequence analysis revealed that the last exon of the short transcripts contains an Open Reading Frame (ORF) and thus putatively can produce peptides. We found ST1C4 expression across different human tissues whereas ST1C5 was mainly restricted to leukocytes, specifically neutrophils. A high up-regulation (10-fold) of short transcripts was observed in isolated human blood neutrophils after activation with lipopolysaccharide. Parallel analyses by liquid chromatography-mass spectrometry identified peptides corresponding to C-terminal region of A1AT in supernatants of activated but not naïve neutrophils. Herein we report for the first time a tissue specific expression and regulation of short transcripts of *SERPINA1* gene, and the presence of C-terminal peptides in supernatants from activated neutrophils, *in vitro*. This gives a novel insight into the studies on the transcription of *SERPINA1* gene.

## Introduction

Gene expression is responsible for the synthesis of functional gene products, typically proteins. The question of how many genes are present in the human genome led to the collection of known protein-coding genes not only in the human genome but also in *Arabidopsis*, worm, fly and mouse [1–4]. The findings revealed that there is a significantly greater amount of

decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

transcriptional output from genomes than anticipated by the collection of annotated protein-coding transcripts [5]. New proteomic high-throughput technologies provide a strong evidence of the existence of non-canonical protein isoforms generated by the translation outside of the annotated protein-coding genes [6,7]. A variety of short open reading frames (ORFs), long noncoding RNAs and pseudogenes are coding short transcripts [8,9], which may express unrecognized biological functions.

The serpins (serine proteinase inhibitors) are proteins found in plants, animals, and viruses [10–12]. These proteins participate in the regulation of the coagulation, fibrinolysis, complement activation, angiogenesis, apoptosis, inflammation, neoplasia, viral pathogenesis, among others [13]. Human genome encodes 16 serpin clades, termed A to P. So far, 36 human serpins have been identified, among which 29 have a primary function in the regulation of the proteolytic activity. Mutations in serpin genes are associated with different human diseases, globally named serpinopathies [14]. Alternative splicing results in multiple transcript variants of serpin genes, however, to our knowledge; current databases contain no information about noncoding transcripts in human serpins.

Human *SERPINA1* gene is located on chromosome 14q31-32.3 [15], which covers approximately 12.2 kb, and has four coding exons, three untranslated exons and six introns. *SERPINA1* gene is expressed and translated in different tissues, and encodes alpha1-antitrypsin (A1AT), a prototypical member of the serpin superfamily of proteins. The regulation of A1AT expression is controlled by different promoters and transcription start sites located in the 5'UTR of the gene [16–18]. In hepatocytes, transcription typically begins in exon 1C [17,19] whereas in monocytes and macrophages A1AT transcribed mainly from exons 1A and 1B [16,18,20]. One particular feature of *SERPINA1* gene is its high transcriptional complexity i.e. according to the Emsembl Genome Browser, human *SERPINA1* gene has 19 transcripts (splice variants), generated in a stimulus- and cell-type specific manner [21–23]. Different types of alternative splicing mechanisms generate this complexity of transcripts, including exon skipping and alternative 5' and 3' splice site usage. Most of the *SERPINA1* transcript variants are produced by the alternative splicing between the non-coding exons 1A, 1B and 1C.

Remarkably, a recent study on ovine *SERPINA1* gene expression found short transcripts lacking exons II or III [24]. Because the *SERPINA1* gene exon and intron organization is similar among the species, theoretically short transcripts can also be present in human *SERPINA1* gene. In support of our prediction, herein we identified novel short transcripts of *SERPINA1* gene and provided evidence for tissue specific expression and regulation of these transcripts.

## Materials and Methods

### RNA extraction, cDNA synthesis and *SERPINA1* expression analysis by RT-PCR

Peripheral blood samples were obtained from subjects previously tested for A1ATD, and informed consent for this study was signed. The study was approved by the ethics committee of Instituto de Salud Carlos III. RNA isolation from peripheral blood leukocytes was performed using RNAeasy kit (Qiagen) following manufacturer's recommendations. The cDNA synthesis was generated using the Maxima First Strand cDNA Synthesis kit (Thermo Scientific). To amplify *SERPINA1* transcripts, primers located in exon 1C: 1C\_F:5' ctgtctcctc agcttcaggc3' and 1A: 1A\_F:5' tgaggagagcaggaaaggaca3' were used in combination with a reverse primer located in exon V, EX5\_R:5' ccatgaagaggggagacttgg3'. The PCR was performed under the following conditions: 35 cycles of 94°C for 45s, 60°C for 45s, and 72°C for 45s. Amplified products were visualized in 1–2% agarose gels. PCR products were purified by using PCR purification Kit (Qiagen) and subsequently cloned into pGEM-T

easy vector (Promega). Ligation reactions were used to transform DH5 $\alpha$  competent cells. Clones containing PCR products were selected by blue/white colony screening and standard ampicillin selection. Positive transformants were analyzed by the PCR and sequenced using *SERPINA1* primers.

### Quantification of ST1C4 and ST1C5 by QT-PCR

The expression levels of two short transcripts, ST1C4 and ST1C5, were independently analyzed by real time quantitative PCR (qPCR). cDNA samples were diluted and amplified using Taqman Fast advance master mix in a 7500 Fast Real Time PCR System (Applied Biosystems) with recommended PCR cycle conditions. To amplify the ST1C4 or ST1C5 transcripts two different reactions were performed using boundary-spanning primers for the sequence covering the 1C-exon4 junction ST1C4\_F:5' tgggacagtgaatc/gtctgc3' or the spanning region between 1C-exon5, ST1C5v3\_F:5' acctgggacagtgaatc/gccgtg3' in combination with a reverse primer in exon 5 for both reactions EX5\_R:5' ccatgaagaggggagacttg3'. Specific fluorescent-labelled Taqman probes were selected from the Universal probe library, UPL, Roche. UPL probe #57 was used for the ST1C4 transcript and the probe #3 was selected for the ST1C5 transcript. GUSB gene was used as an endogenous control to normalize the expression levels of the transcripts. All experiments were performed in triplicate and data were analyzed by Applied Biosystems 7500 software v2.0. Relative expression was calculated by using the comparative Ct method and obtaining the fold-change value ( $\Delta\Delta Ct$ ).

We also analyzed *SERPINA1* full transcripts and compared these with expression of identified short transcripts. We individually analyzed alternative full transcripts generated by the use of exons 1A, 1B and 1C by using primers and methods reported before [25].

### ST1C4 and ST1C5 expression in cDNA samples from human tissues

We used a set of first-strand cDNAs from different tissues for analysis of short transcripts ST1C4 and ST1C5. We analyzed 16 commercially available first-strand cDNAs from 16 different human tissues (Human MTC Panel I and II, Clontech) as described before [25]. We performed at least two experiments in triplicate for all tissues. We normalized the expression of short transcript to the expression obtained in a leukocyte sample.

### PBMCs and neutrophil isolation from human blood

Human Peripheral blood mononuclear cells (PBMCs) were isolated from blood of healthy volunteers using Lymphosep discontinuous gradient centrifugation. PBMCs were re-suspended at a density of  $5 \times 10^6$  cells/ml in RPMI-1640 with 2 mM N-acetyl-L-alanyl-L-glutamine (Gibco, Life Technologies) supplemented with 1% nonessential amino acids, 2% sodium pyruvate and 20 mM HEPES and plated. We incubated cells for 75 min at 37°C and 5% CO<sub>2</sub> to allow monocytes to adhere to the cell culture plates. Afterwards, we removed non-adherent cells by washing with PBS containing Mg<sup>2+</sup> and Ca<sup>2+</sup> (Gibco, Life Technologies) and fresh medium was added. Next day, adherent PBMCs were used for experiments.

Neutrophils were isolated from the freshly obtained peripheral blood of healthy donors using Polymorphoprep<sup>TM</sup> (Axis-Shield PoC AS, Oslo, Norway) according to the manufacturers' recommendations. The neutrophil purity was typically 98% as determined by cytopspins and cell viability was more than 90% according to trypan blue staining. Before starting the experiment, neutrophils ( $3 \times 10^6$  per well) were always left alone for 30–40 min in cell culture plates pre-coated with fetal calf serum.

Freshly isolated neutrophils were treated with 10 ng/ml LPS for 5h. In parallel, cells were incubated with appropriate buffer.

## Sample preparation and LC-MS analysis

C-10, C-21 and C-36 standard peptides (JPT, Innovative peptide solutions, Berlin, Germany) were dissolved in RPMI Medium, to obtain a final concentration of 20  $\mu\text{M}$ . All supernatants from neutrophil cultures (controls and LPS-stimulated) were filtered using Microcon<sup>®</sup> centrifugal filters (Millipore Corporation, Bedford, MA, USA) with 10 kDa molecular weight cut-off (MWCO), to remove any large protein and concentrate the C-terminal peptides. 100  $\mu\text{L}$  aliquots of both standard solutions and real samples were acidified by addition of 2  $\mu\text{L}$  of formic acid (FA), to reach a final pH of about 3.0 was reached, and analyzed by LC-MS.

The LC-MS system (Thermo Finnigan, San Jose, CA, USA) consisted of a thermostated column oven Surveyor autosampler controlled at 25°C; a quaternary gradient Surveyor MS pump; a diode array detector and a Linear Trap Quadrupole (LTQ) mass spectrometer with electrospray ionization ion source controlled by Xcalibur software 2.0.7. Analytes were separated by RP-HPLC on a Jupiter (Phenomenex, Torrance, CA, USA) C<sub>18</sub> column (150 x 2 mm, 4  $\mu\text{m}$ , 90 Å particle size) using a linear gradient (2–60% solvent B in 60 min) in which solvent A consisted of 0.1% aqueous FA and solvent B of acetonitrile (ACN) containing 0.1% FA. Flow-rate was 0.2 mL/min. Mass spectra were generated in positive ion mode under constant instrumental conditions: source voltage 5.0 kV, capillary voltage 46 V, sheath gas flow 40 (arbitrary units), auxiliary gas flow 10 (arbitrary units), sweep gas flow 1 (arbitrary units), capillary temperature 200°C, tube lens voltage –105 V. MS/MS spectra, obtained by CID studies in the linear ion trap, were performed with an isolation width of 3 Th  $m/z$ , the activation amplitude was 35% of ejection RF amplitude that corresponds to 1.58 V. All analyses were performed in triplicate.

## Statistical analysis

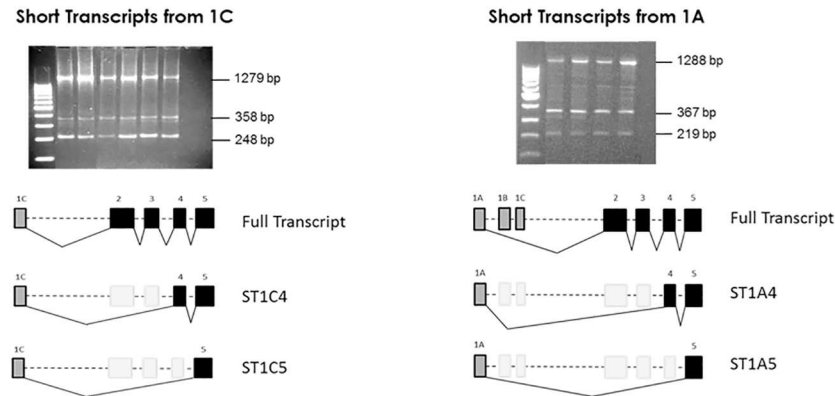
Statistical Package (SPSS for Windows, release 21.0) was used for the calculations. An independent two-sample t-test was used. The differences in the means of cell experimental results were analyzed for their statistical significance using one-way ANOVA combined with a multiple-comparison procedure (Scheffe multiple range test), with an overall significance level of  $p = 0.05$ .

## Results and Discussion

### Identification of *SERPINA1* short transcripts

Peripheral blood leukocytes (e.g., T cells, monocytes, neutrophils) are readily accessible source of information about the health and physiological state of an individual. By comparing gene expression profiles of leukocytes, it is possible to identify genes with cell-type-specific expression patterns [26]. To analyze *SERPINA1* expression we prepared cDNA from human peripheral blood leukocytes and used a primer pair for a full-length *SERPINA1* cDNA (1C\_F and EX5\_R) to amplify *SERPINA1* transcripts. When agarose gels visualized amplified products, we identified an expected 1279 bp band corresponding to the full-length A1AT. Cloning and sequence analyses confirmed the presence of the exon 1C and four coding exons (exon II to exon V) encoding the 419 aa A1AT protein. This transcript corresponded to the Ensembl transcript ENST00000393087.

Surprisingly, in addition to the 1279 bp band, we constantly observed two bands of few hundred base pairs, one of which (a band of 358 bp) revealed a short transcript of A1AT (Fig 1). After cloning, several colonies were obtained and sequencing showed a transcript with exon 1C directly spliced to exon IV (ST1C4), with an exon structure 1C-IV-V, lacking exons II and III. A shortest band of 248 bp revealed another transcript of A1AT with exon 1C joined to



**Fig 1. Expression of A1AT transcripts in peripheral blood cells.** Gel electrophoresis shows cDNA transcript variants of *SERPINA1* gene generated from exon 1C or exon 1A. Three different transcripts were obtained by PCR using RNA extracted from whole blood samples. The larger bands in both cases corresponded to the full length *SERPINA1* transcript. The other lower bands represent short transcripts of the gene which included the untranslated exons 1C or 1A joined to exon IV and V (ST1C4 and ST1A4) or joined to exon V (ST1C5 and ST1A5).

doi:10.1371/journal.pone.0170533.g001

exon V (ST1C5) missing exons II, III and IV. Sequence information of these new short transcripts is available through the following NCBI GenBank accession numbers: KU535895 (ST1C4) and KU535894 (ST1C5).

Since *SERPINA1* alternative splicing events typically occur among the first non-coding exons, we investigated whether the newly identified short transcripts can possibly originate from exon 1A. Using a primer pair located in exon 1A (1A\_F) and exon V (EX5\_R), three main bands and few less evident bands occurred, similarly as we found by using the primer at exon 1C (1C\_F) (Fig 1). The band of 1288 bp represented the full *SERPINA1* transcript starting at exon 1A joined to the rest of coding exons II, III, IV and V. This transcript corresponded to the Ensembl transcript ENST00000355814.

Several colonies expressing an intermediate band of 367 bp were obtained, and sequencing analysis revealed a transcript structure containing exon 1A directly bound to exons IV and V (ST1A4) (Fig 1). Finally, a shortest band of 219 bp was also cloned and sequencing analysis revealed a short transcript with exon 1A directly joined to exon V, lacking all other coding exons (ST1A5). Sequence information of *SERPINA1* short transcripts transcribed from exon 1A can be accessed through NCBI GenBank accession numbers: KU755451 for the ST1A4 short transcript and KU755450 for the ST1A5 short transcript. Altogether, our results demonstrate that short transcripts can be transcribed from either exon 1C or 1A.

A previous study on ovine *SERPINA1* gene expression also found short transcripts lacking exons II or III [24]. However, human genomic databases (Ensembl or UCSC genome browser) do not contain information regarding short transcripts of *SERPINA1* gene identified in this study. Importantly, however, we found a RNA-Seq data set [27], presenting expression and alternative splicing analysis of blood leukocytes derived from controls and Parkinson's patients. In this data set, we found evidence supporting the presence of *SERPINA1* short transcripts. By looking at single reads that map to different exon/exon junctions, we identified several reads, in which the exon 1A binds directly to exon IV, lacking the exons II and III, similar as occurred with the transcript ST1A4 described here.

Interestingly, *in silico* analysis showed that, ST1C4 and ST1C5 as well as ST1A4 and ST1A5 have an ORF in exon V starting at Met351 of the full-length A1AT protein. Further analysis confirmed that in the exon V, in addition to Met351, there are three more methionine

Short transcript, ST1C5

CAGCCTCCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGACGAGGACAGGGCCCT  
 GTCTCCTCAGCTTCAGGCACCCACTGACCTGGGACAGTGAATCGCCGTGCATAAGGCT  
 GTGCTGACCATCGACGAGAAAAGGGACTGAAGCTGCTGGGGCCATGTTTTAGAGGCCATA  
 CCCATGTCTATCCCCCGAGGTCAAGTCAACAAACCCCTTGTCTTCTAATGATTGAACAA  
 ATACCAAGTCTCCCCTCTCATGGAAAAGTGGTGAATCCACCCAAAAATAA

PEPTIDE 1 SEQUENCE 44 AA  
**MFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK**  
 PEPTIDE 2 SEQUENCE 37 AA  
**MSIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK**  
 PEPTIDE 3 SEQUENCE 21 AA  
**MIEQNTKSPLFMGKVVNPTQK**  
 PEPTIDE 4 SEQUENCE 10 AA  
**MGKVVNPTQK**

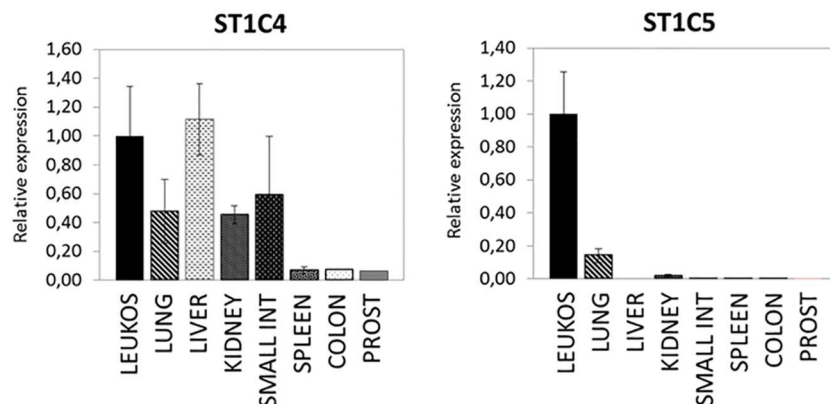
**Fig 2. Sequence of the short transcript ST1C5.** The exon 5 sequence has an ORF with several in frame ATG codons (these are underlined). Predicted amino acid sequences of potential peptides of 44, 36, 21 and 10 amino acids, generated by these ORF are shown.

doi:10.1371/journal.pone.0170533.g002

residues. Hence, putative peptides generated by the short transcripts would have the same C-terminal amino acid sequence as a full-length A1AT protein, i.e. the last 44aa, 36aa, 21aa or 10aa, respectively (Fig 2).

Differential expression and induction of short ST1C4 and ST1C5 transcripts in blood monocytes and neutrophils by LPS

To investigate, if short transcripts of *SERPINA1* gene have a tissue specific expression, we performed a quantitative analysis focusing on ST1C4 and ST1C5, using cDNA from various human tissues (Fig 3). These cDNA panels represent pooled tissue samples obtained from different individuals, and allow accurate and precise gene expression analysis in a wide variety of tissues. The expression profile of the ST1C4 transcript resembled that known for A1AT, i.e. showed a high expression in the liver, lung, kidney, small intestine and leukocytes [28,29]. In contrast, the expression of ST1C5 transcripts was almost exclusively restricted to leukocytes. As illustrated in Fig 3, ST1C5 was detected in the lung tissue but expression levels accounted



**Fig 3. Expression profiles in different tissue samples.** Tissue expression of short transcripts ST1C4 and ST1C5. Short transcripts show differential expression patterns in human tissues. Data were normalized to the expression detected in leukocytes. Histograms represent mean expression values of two experiments performed in triplicates. Bars represent mean (SD). LEUKO: Leukocytes, SMALL INT: Small Intestine, PROST: Prostate.

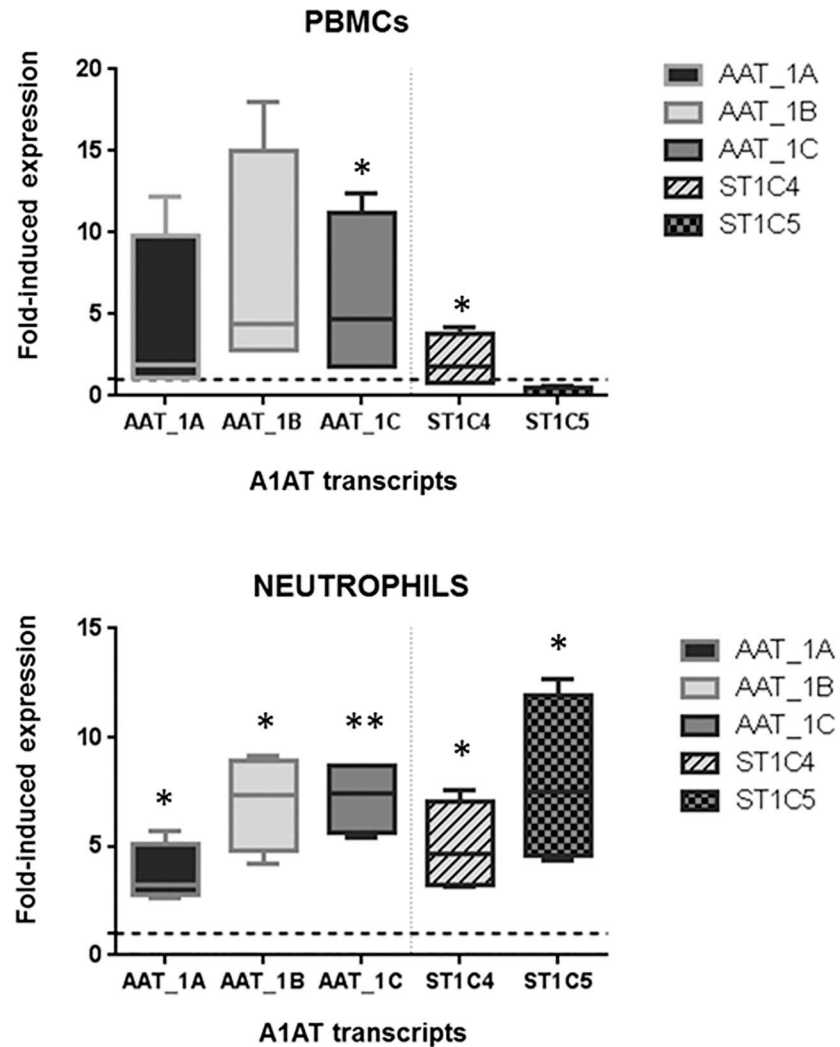
doi:10.1371/journal.pone.0170533.g003

only for 20% of those seen in leukocytes. Although the commercial cDNA panels used for expression patterning of short transcripts provided a nice overview about their expression, it would be useful to perform a more detailed analysis in future studies on specific tissues and cell types.

Since leukocytes showed the highest expression of both ST1C4 and ST1C5 transcripts, we next investigated the expression and regulation of these transcripts in human blood isolated PBMCs and neutrophils. Cytokines, chemokines, hormones and other substances control the constitutive and modulated expression, and release of A1AT [17,30]. These factors act specifically on different cell types and probably control the production of A1AT at sites of inflammation. Bacterial lipopolysaccharide (LPS), an inflammatory activator, modulates synthesis of A1AT in different cell types including neutrophils, monocytes and macrophages [31]. Therefore, we prepared freshly isolated human blood PBMCs and neutrophil cultures that were stimulated or not with LPS. As expected, LPS increased the expression of the full A1AT transcript in PBMCs and neutrophils (Fig 4). In PBMCs, LPS also induced the ST1C4 expression by 2 to 4-fold, but had no effects on ST1C5 expression. Neutrophils stimulated with LPS showed increased expression of both ST1C4 (by 2 to 8-fold) and ST1C5 (by 5 to 10-fold) if compared to non-treated neutrophils (Fig 4). In contrast, T lymphocytes did not show significant expression of either A1AT or short transcripts (data not shown).

### Detection of A1AT short C-terminal peptide in LPS-stimulated neutrophils

Since A1AT short transcripts ST1C4 and ST1C5 were highly induced in LPS-stimulated neutrophils, we speculated about the possibility to find A1AT peptides in cell culture supernatants. Therefore, we next performed LC-MS analysis in the cell culture supernatants from LPS-stimulated and not stimulated neutrophils. Fig 5A shows the MS profile obtained from non-stimulated cells (used as “controls”) and Fig 5B and 5C illustrate the profiles of LPS-stimulated cells (for 2 h), obtained from two independent experiments (indicated as LPS1 and LPS2). While non-stimulated cells (Fig 4A) do not show any significant signal, thus representing essentially a “blank”, LPS-stimulated cells (Fig 5B and 5C) contain clear signatures (at 690,25; 827,99; 1034,32 and 1378,88 m/z values) corresponding to the multiple-charged ion ( $[M + H]^{6+}$ ;  $[M + H]^{5+}$ ;  $[M + H]^{4+}$ ;  $[M + H]^{3+}$ , respectively) with  $4134,9 \pm 1$  Da molecular size. Using the Swissprot database, this fragment unambiguously was identified as a C-terminal fragment of A1AT (4135 Da). To “validate” this, in both experiments the real sample was spiked with an aliquot (1  $\mu$ M) of the standard, synthetic C-36 fragment. Not surprisingly, the pattern of m/z signatures was identical to that described above, the only difference being observed in the intensity of multiple-charged ions. Spiking the “blank” with 1  $\mu$ M standard C-36 fragment resulted in the generation of a MS profile that was similar to that shown in Fig 5B and 5C. These latter results are presented in the inset of Fig 5A. The LC-MS runs performed on supernatants of cells that were LPS-stimulated for 5 h in two independent experiments gave essentially the same results to those illustrated in Fig 5 (data not shown). The results from the LC-MS experiments for the first time identify C-terminal peptide of A1AT in the supernatants of activated neutrophils but unfortunately do not answer the question of whether this peptide is generated after endogenous A1AT cleavage or is newly produced and secreted. Cleavage by target proteases (like neutrophil elastase or proteinase 3) triggers a conformational change of A1AT, in which the N-terminal portion of the reactive center loop spontaneously inserts into  $\beta$ -sheet A, carrying the bound protease (covalently linked as an acyl-enzyme intermediate) along with it. Cleaved C-terminal part of the A1AT typically remains bound to the larger part of the cleaved A1AT. However, under the



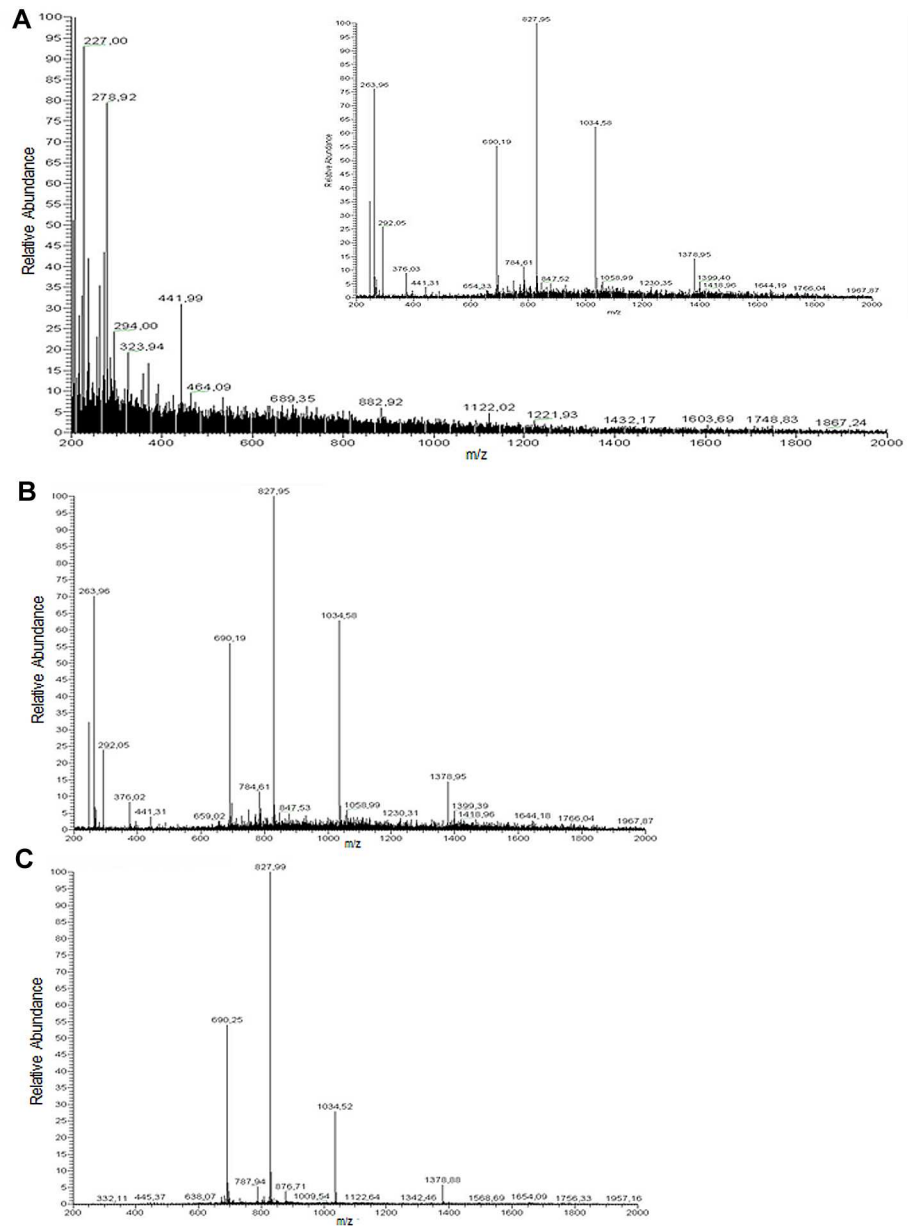
**Fig 4. Induction of *SERPINA1* transcripts by LPS-treatment in isolated human blood PBMCs and neutrophils.** Treatment of PBMCs with 100 ng LPS (n = 4 independent experiments) and neutrophils with 10ng LPS for 5 hours (n = 4 independent experiments) resulted in enhanced expression of full transcripts (1A, 1B and 1C) and short transcripts ST1C4 and ST1C5 relatively to the basal expression in untreated cells (represented by a dashed line). Box plots display variation in samples. Statistically significant values are shown by \* (p<0.05) or \*\* (p<0.01).

doi:10.1371/journal.pone.0170533.g004

specific conditions free C-terminal peptide of A1AT can be generated. For example, Frochoux V et al. [32] shows that the reaction of placental A1AT with HTRA1 leads not only to the formation of a protease-protease inhibitor complex, but also to the formation of cleavage fragments of A1AT. Therefore, future studies are needed to specifically address the question of whether the free C-terminal peptides of A1AT are newly synthesized.

So far, all described alternative human *SERPINA1* transcripts differ regarding their sites of transcription initiation and alternative splicing events, located in exons 1A, 1B and 1C at the 5'UTR region. Intriguingly, the new short transcripts variants are generated by alternative splicing affecting the coding exons, and all short transcripts include the last exon V of the gene. Exon V encodes a very important functional part of the A1AT protein, since it includes the active site (Met358-Ser359) as part of the reactive site loop of the A1AT, which acts as a





**Fig 5. MS profiles of supernatants from cell cultures.** A) shows the MS spectrum of a control sample ("blank") and of the same sample spiked with an aliquot (1 μM) of the standard C-36 fragment (inset). B and C) show to the profiles of LPS-stimulated cells for 2 h, obtained from two different experiments (indicated as LPS1 and LPS2). The multiple-charged ions ( $[M + H]^6+$ ;  $[M + H]^5+$ ;  $[M + H]^4+$ ;  $[M + H]^3+$ ), represent the fragment with  $4134,9 \pm 1$  Da molecular size. For further details, see the text.

doi:10.1371/journal.pone.0170533.g005

docking site for binding and neutralizing neutrophil elastase [33]. Therefore, the 44 aa peptides are expected to contain the active site region and protease inhibitory activity. In support, the 44-residue C-terminal fragment of A1AT detected in placenta and human tissues [34], was found to inhibit neutrophil and pancreatic elastases [35].

A1AT is a well-known protease inhibitor forming stable complexes with a range of serine proteases, specifically neutrophil elastase and proteinase 3. A1AT can be targeted by non-specific proteases (e.g., matrix metalloproteases) without the formation of stable inhibitor

complexes. (MEROPS database, <http://merops.sanger.ac.uk/index.shtml>). In both cases, interaction between A1AT and proteases results in cleavage of A1AT and irreversible loss of its inhibitory activity. Different studies provide evidence that native A1AT is proteolytically processed into firmly attached larger N-terminal and shorter C-terminal fragments [36]. Interestingly, in contrast to the biochemical and structural data, variants of A1AT C-fragments have been identified in human tissues and body fluids [34,37–41]. The physiological functions for these peptides have also been reported, like NK-cell suppression (CRISPP peptide)[42], extracellular matrix protection (SPAAT peptide) [43], pro- or anti-inflammatory immune modulating functions [41,44], and inhibition of HIV entry (VIRIP peptide) [40]. We previously reported that the C-terminal fragment of A1AT forms amyloid-like fibrils, activates human monocytes and is present in atherosclerotic plaques [45]. We also found C-terminal peptide of A1AT in the lungs of patients with chronic obstructive pulmonary diseases. [41]. Recent studies revealed that so called cell-penetrating peptides (CPPs) have the ability to cross cellular membranes, either alone or in association with bioactive cargo. The C-terminal peptide of A1AT, 105Y (CSIPPEVKFNKPFVYLI), has been characterized as endoport for siRNA [46]. Hence, the effects of C-terminal peptides *in vivo* are likely to be complex. It remains to be confirmed if free C-terminal peptide(s) of A1AT are exclusively produced from the A1AT cleavage or/and can be *de novo* synthesized and released by different cells.

In summary, we identified novel short transcripts of the *SERPINA1* gene, differentially expressed in tissues and cells. We also provide experimental evidence that the expression of short transcripts can be regulated and that C-terminal peptides of A1AT can be found in free form in LPS stimulated neutrophils.

## Acknowledgments

We thank collaborators from the REDAAT (Spanish Registry of Alpha-1 Antitrypsin deficiency patients) and all members of the Human Genetics Area of the ISCIII for their support.

## Author Contributions

**Conceptualization:** BMD SJ JS.

**Formal analysis:** NM NA GG IC.

**Funding acquisition:** BMD.

**Investigation:** NM NA PI MF GG BL MTM IC.

**Methodology:** NM SJ PI MF BMD.

**Project administration:** JS SJ BMD.

**Resources:** SJ BL MTM IC.

**Software:** IC.

**Supervision:** SJ BMD.

**Validation:** NM NA PI MF GG SJ.

**Visualization:** NM PI JS SJ BMD.

**Writing – original draft:** NM SJ JS BMD.

**Writing – review & editing:** JS SJ BMD.

## References

1. Carninci P (2006) Tagging mammalian transcription complexity. *Trends Genet* 22: 501–510. doi: [10.1016/j.tig.2006.07.003](https://doi.org/10.1016/j.tig.2006.07.003) PMID: [16859803](https://pubmed.ncbi.nlm.nih.gov/16859803/)
2. Johnson JM, Edwards S, Shoemaker D, Schadt EE (2005) Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. *Trends Genet* 21: 93–102. doi: [10.1016/j.tig.2004.12.009](https://doi.org/10.1016/j.tig.2004.12.009) PMID: [15661355](https://pubmed.ncbi.nlm.nih.gov/15661355/)
3. Kapranov P, Willingham AT, Gingeras TR (2007) Genome-wide transcription and the implications for genomic organization. *Nat Rev Genet* 8: 413–423. doi: [10.1038/nrg2083](https://doi.org/10.1038/nrg2083) PMID: [17486121](https://pubmed.ncbi.nlm.nih.gov/17486121/)
4. Willingham AT, Gingeras TR (2006) TUF love for "junk" DNA. *Cell* 125: 1215–1220. doi: [10.1016/j.cell.2006.06.009](https://doi.org/10.1016/j.cell.2006.06.009) PMID: [16814704](https://pubmed.ncbi.nlm.nih.gov/16814704/)
5. Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SP et al. (2002) Large-scale transcriptional activity in chromosomes 21 and 22. *Science* 296: 916–919. doi: [10.1126/science.1068597](https://doi.org/10.1126/science.1068597) PMID: [11988577](https://pubmed.ncbi.nlm.nih.gov/11988577/)
6. Ingolia NT, Brar GA, Stern-Ginossar N, Harris MS, Talhouarne GJ, Jackson SE et al. (2014) Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep* 8: 1365–1379. doi: [10.1016/j.celrep.2014.07.045](https://doi.org/10.1016/j.celrep.2014.07.045) PMID: [25159147](https://pubmed.ncbi.nlm.nih.gov/25159147/)
7. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R et al. (2014) A draft map of the human proteome. *Nature* 509: 575–581. doi: [10.1038/nature13302](https://doi.org/10.1038/nature13302) PMID: [24870542](https://pubmed.ncbi.nlm.nih.gov/24870542/)
8. Galindo MI, Pueyo JI, Fouix S, Bishop SA, Couso JP (2007) Peptides encoded by short ORFs control development and define a new eukaryotic gene family. *PLoS Biol* 5: e106. doi: [10.1371/journal.pbio.0050106](https://doi.org/10.1371/journal.pbio.0050106) PMID: [17439302](https://pubmed.ncbi.nlm.nih.gov/17439302/)
9. Laumont CM, Daouda T, Laverdure JP, Bonneil E, Caron-Lizotte O, Hardy MP et al. (2016) Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat Commun* 7: 10238. doi: [10.1038/ncomms10238](https://doi.org/10.1038/ncomms10238) PMID: [26728094](https://pubmed.ncbi.nlm.nih.gov/26728094/)
10. Grosse-Holz FM, van der Hoorn RA (2016) Juggling jobs: roles and mechanisms of multifunctional protease inhibitors in plants. *New Phytol* 210: 794–807. doi: [10.1111/nph.13839](https://doi.org/10.1111/nph.13839) PMID: [26800491](https://pubmed.ncbi.nlm.nih.gov/26800491/)
11. Heit C, Jackson BC, McAndrews M, Wright MW, Thompson DC, Silverman GA et al. (2013) Update of the human and mouse SERPIN gene superfamily. *Hum Genomics* 7: 22. doi: [10.1186/1479-7364-7-22](https://doi.org/10.1186/1479-7364-7-22) PMID: [24172014](https://pubmed.ncbi.nlm.nih.gov/24172014/)
12. Zheng D, Chen H, Bartee MY, Williams J, Davids JA, Huang E et al. (2012) Virus-derived anti-inflammatory proteins: potential therapeutics for cancer. *Trends Mol Med* 18: 304–310. doi: [10.1016/j.molmed.2012.03.006](https://doi.org/10.1016/j.molmed.2012.03.006) PMID: [22554906](https://pubmed.ncbi.nlm.nih.gov/22554906/)
13. van GD, Sharp P, Morgan K, Kalsheker N (2003) Serpins: structure, function and molecular evolution. *Int J Biochem Cell Biol* 35: 1536–1547. PMID: [12824063](https://pubmed.ncbi.nlm.nih.gov/12824063/)
14. Roussel BD, Irving JA, Ekeowa UI, Belorgey D, Haq I, Ordonez A et al. (2011) Unravelling the twists and turns of the serpinopathies. *FEBS J* 278: 3859–3867. doi: [10.1111/j.1742-4658.2011.08201.x](https://doi.org/10.1111/j.1742-4658.2011.08201.x) PMID: [21624056](https://pubmed.ncbi.nlm.nih.gov/21624056/)
15. Billingsley GD, Walter MA, Hammond GL, Cox DW (1993) Physical mapping of four serpin genes: alpha 1-antitrypsin, alpha 1-antichymotrypsin, corticosteroid-binding globulin, and protein C inhibitor, within a 280-kb region on chromosome 14q32.1. *Am J Hum Genet* 52: 343–353. PMID: [8381582](https://pubmed.ncbi.nlm.nih.gov/8381582/)
16. Hafeez W, Ciliberto G, Perlmutter DH (1992) Constitutive and modulated expression of the human alpha 1 antitrypsin gene. Different transcriptional initiation sites used in three different cell types. *J Clin Invest* 89: 1214–1222. doi: [10.1172/JCI115705](https://doi.org/10.1172/JCI115705) PMID: [1556183](https://pubmed.ncbi.nlm.nih.gov/1556183/)
17. Kalsheker N, Morley S, Morgan K (2002) Gene regulation of the serine proteinase inhibitors alpha1-antitrypsin and alpha1-antichymotrypsin. *Biochem Soc Trans* 30: 93–98. PMID: [12023832](https://pubmed.ncbi.nlm.nih.gov/12023832/)
18. Perlino E, Cortese R, Ciliberto G (1987) The human alpha 1-antitrypsin gene is transcribed from two different promoters in macrophages and hepatocytes. *EMBO J* 6: 2767–2771. PMID: [3500042](https://pubmed.ncbi.nlm.nih.gov/3500042/)
19. Ciliberto G, Dente L, Cortese R (1985) Cell-specific expression of a transfected human alpha 1-antitrypsin gene. *Cell* 41: 531–540. PMID: [2985281](https://pubmed.ncbi.nlm.nih.gov/2985281/)
20. Rollini P, Fournier RE (2000) Differential regulation of gene activity and chromatin structure within the human serpin gene cluster at 14q32.1 in macrophage microcell hybrids. *Nucleic Acids Res* 28: 1767–1777. PMID: [10734196](https://pubmed.ncbi.nlm.nih.gov/10734196/)
21. Clemmensen SN, Jacobsen LC, Rorvig S, Askaa B, Christenson K, Iversen M et al. (2011) Alpha-1-antitrypsin is produced by human neutrophil granulocytes and their precursors and liberated during granulocyte exocytosis. *Eur J Haematol* 86: 517–530. doi: [10.1111/j.1600-0609.2011.01601.x](https://doi.org/10.1111/j.1600-0609.2011.01601.x) PMID: [21477074](https://pubmed.ncbi.nlm.nih.gov/21477074/)
22. Van't Wout EF, van SA, Lomas DA, Stolk J, Marciniak SJ, Hiemstra PS (2015) Function of monocytes and monocyte-derived macrophages in alpha1-antitrypsin deficiency. *Eur Respir J* 45: 365–376. doi: [10.1183/09031936.00046114](https://doi.org/10.1183/09031936.00046114) PMID: [25323228](https://pubmed.ncbi.nlm.nih.gov/25323228/)

23. Wood AM, Simmonds MJ, Bayley DL, Newby PR, Gough SC, Stockley RA (2008) The TNFalpha gene relates to clinical phenotype in alpha-1-antitrypsin deficiency. *Respir Res* 9: 52. doi: [10.1186/1465-9921-9-52](https://doi.org/10.1186/1465-9921-9-52) PMID: [18620570](https://pubmed.ncbi.nlm.nih.gov/18620570/)
24. Marchitelli C, Crisa A, Mostarda E, Napolitano F, Muioli B (2013) Splicing variants of SERPINA1 gene in ovine milk: characterization of cDNA and identification of polymorphisms. *PLoS One* 8: e73020. doi: [10.1371/journal.pone.0073020](https://doi.org/10.1371/journal.pone.0073020) PMID: [24009725](https://pubmed.ncbi.nlm.nih.gov/24009725/)
25. Matamala N, Martinez MT, Lara B, Perez L, Vazquez I, Jimenez A et al. (2015) Alternative transcripts of the SERPINA1 gene in alpha-1 antitrypsin deficiency. *J Transl Med* 13: 211. doi: [10.1186/s12967-015-0585-y](https://doi.org/10.1186/s12967-015-0585-y) PMID: [26141700](https://pubmed.ncbi.nlm.nih.gov/26141700/)
26. Shen-Orr SS, Tibshirani R, Khatri P, Bodian DL, Staedtler F, Perry NM et al. (2010) Cell type-specific gene expression differences in complex tissues. *Nat Methods* 7: 287–289. doi: [10.1038/nmeth.1439](https://doi.org/10.1038/nmeth.1439) PMID: [20208531](https://pubmed.ncbi.nlm.nih.gov/20208531/)
27. Soreq L, Salomonis N, Guffanti A, Bergman H, Israel Z, Soreq H (2015) Whole transcriptome RNA sequencing data from blood leukocytes derived from Parkinson's disease patients prior to and following deep brain stimulation treatment. *Genom Data* 3: 57–60. doi: [10.1016/j.gdata.2014.11.009](https://doi.org/10.1016/j.gdata.2014.11.009) PMID: [26484149](https://pubmed.ncbi.nlm.nih.gov/26484149/)
28. Carlson JA, Rogers BB, Sifers RN, Hawkins HK, Finegold MJ, Woo SL (1988) Multiple tissues express alpha 1-antitrypsin in transgenic mice and man. *J Clin Invest* 82: 26–36. doi: [10.1172/JCI113580](https://doi.org/10.1172/JCI113580) PMID: [3260605](https://pubmed.ncbi.nlm.nih.gov/3260605/)
29. Kelsey GD, Povey S, Bygrave AE, Lovell-Badge RH (1987) Species- and tissue-specific expression of human alpha 1-antitrypsin in transgenic mice. *Genes Dev* 1: 161–171. PMID: [3500093](https://pubmed.ncbi.nlm.nih.gov/3500093/)
30. Yuan ZA, Soprano KJ, Kueppers F (1992) Alpha-1 antitrypsin response of stimulated alveolar macrophages. *J Cell Biochem* 49: 410–416. doi: [10.1002/jcb.240490411](https://doi.org/10.1002/jcb.240490411) PMID: [1429867](https://pubmed.ncbi.nlm.nih.gov/1429867/)
31. Barbey-Morel C, Pierce JA, Campbell EJ, Perlmutter DH (1987) Lipopolysaccharide modulates the expression of alpha 1 proteinase inhibitor and other serine proteinase inhibitors in human monocytes and macrophages. *J Exp Med* 166: 1041–1054. PMID: [3498786](https://pubmed.ncbi.nlm.nih.gov/3498786/)
32. Frochoux V, Hildebrand D, Talke A, Linscheid MW, Schluter H (2014) Alpha-1-antitrypsin: a novel human high temperature requirement protease A1 (HTRA1) substrate in human placental tissue. *PLoS One* 9: e109483. doi: [10.1371/journal.pone.0109483](https://doi.org/10.1371/journal.pone.0109483) PMID: [25329061](https://pubmed.ncbi.nlm.nih.gov/25329061/)
33. Wilczynska M, Fa M, Ohlsson PI, Ny T (1995) The inhibition mechanism of serpins. Evidence that the mobile reactive center loop is cleaved in the native protease-inhibitor complex. *J Biol Chem* 270: 29652–29655. PMID: [8530349](https://pubmed.ncbi.nlm.nih.gov/8530349/)
34. Niemann MA, Narkates AJ, Miller EJ (1992) Isolation and serine protease inhibitory activity of the 44-residue, C-terminal fragment of alpha 1-antitrypsin from human placenta. *Matrix* 12: 233–241. PMID: [1406456](https://pubmed.ncbi.nlm.nih.gov/1406456/)
35. Niemann MA, Baggott JE, Miller EJ (1997) Inhibition of human serine proteases by SPAAT, the C-terminal 44-residue peptide from alpha1-antitrypsin. *Biochim Biophys Acta* 1340: 123–130. PMID: [9217022](https://pubmed.ncbi.nlm.nih.gov/9217022/)
36. Otlewski J, Jelen F, Zakrzewska M, Oleksy A (2005) The many faces of protease-protein inhibitor interaction. *EMBO J* 24: 1303–1310. doi: [10.1038/sj.emboj.7600611](https://doi.org/10.1038/sj.emboj.7600611) PMID: [15775973](https://pubmed.ncbi.nlm.nih.gov/15775973/)
37. Blaurock N, Schmerler D, Hünninger K, Kurzai O, Ludewig K, Baier M et al. (2016) C-Terminal Alpha-1 Antitrypsin Peptide: A New Sepsis Biomarker with Immunomodulatory Function. *Mediators of Inflammation* 2016. Article ID 6129437, 13 pages, 2016.
38. Gerbod-Giannone MC, Del Castillo-Olivares A, Janciauskiene S, Gil G, Hylemon PB (2002) Suppression of cholesterol 7alpha-hydroxylase transcription and bile acid synthesis by an alpha1-antitrypsin peptide via interaction with alpha 1-fetoprotein transcription factor. *J Biol Chem* 277: 42973–42980. doi: [10.1074/jbc.M205089200](https://doi.org/10.1074/jbc.M205089200) PMID: [12223476](https://pubmed.ncbi.nlm.nih.gov/12223476/)
39. Johansson J, Grondal S, Sjoval J, Jorvall H, Curstedt T (1992) Identification of hydrophobic fragments of alpha 1-antitrypsin and C1 protease inhibitor in human bile, plasma and spleen. *FEBS Lett* 299: 146–148. PMID: [1544487](https://pubmed.ncbi.nlm.nih.gov/1544487/)
40. Kramer HB, Lavender KJ, Qin L, Stacey AR, Liu MK, di GK et al. (2010) Elevation of intact and proteolytic fragments of acute phase proteins constitutes the earliest systemic antiviral response in HIV-1 infection. *PLoS Pathog* 6: e1000893. doi: [10.1371/journal.ppat.1000893](https://doi.org/10.1371/journal.ppat.1000893) PMID: [20463814](https://pubmed.ncbi.nlm.nih.gov/20463814/)
41. Subramanyam D, Glader P, von WK, Burneckiene J, Stevens T, Janciauskiene S (2006) C-36 peptide, a degradation product of alpha1-antitrypsin, modulates human monocyte activation through LPS signaling pathways. *Int J Biochem Cell Biol* 38: 563–575. doi: [10.1016/j.biocel.2005.09.021](https://doi.org/10.1016/j.biocel.2005.09.021) PMID: [16384723](https://pubmed.ncbi.nlm.nih.gov/16384723/)
42. Cercek L, Cercek B (1993) Cancer-associated SCM-recognition, immunodefense suppression, and serine protease protection peptide. Part III. CRISPP peptide protection of serine proteases against inhibition. *Cancer Detect Prev* 17: 447–454. PMID: [7691406](https://pubmed.ncbi.nlm.nih.gov/7691406/)

43. Niemann MA, Baggott JE, Miller EJ (1997) Binding of SPAAT, the 44-residue C-terminal peptide of alpha 1-antitrypsin, to proteins of the extracellular matrix. *J Cell Biochem* 66: 346–357. PMID: [9257191](#)
44. Zelvyte I, Stevens T, Westin U, Janciauskiene S (2004) alpha1-antitrypsin and its C-terminal fragment attenuate effects of degranulated neutrophil-conditioned medium on lung cancer HCC cells, in vitro. *Cancer Cell Int* 4: 7. doi: [10.1186/1475-2867-4-7](#) PMID: [15555067](#)
45. Dichtl W, Moraga F, Ares MP, Crisby M, Nilsson J, Lindgren S et al. (2000) The carboxyl-terminal fragment of alpha 1-antitrypsin is present in atherosclerotic plaques and regulates inflammatory transcription factors in primary human monocytes. *Mol Cell Biol Res Commun* 4: 50–61. doi: [10.1006/mcbr.2000.0256](#) PMID: [11152628](#)
46. Bartz R, Fan H, Zhang J, Innocent N, Cherrin C, Beck SC et al. (2011) Effective siRNA delivery and target mRNA degradation using an amphipathic peptide to facilitate pH-dependent endosomal escape. *Biochem J* 435: 475–487. doi: [10.1042/BJ20101021](#) PMID: [21265735](#)