The $\alpha_1$ antitrypsin gene and chronic lung disease

N Kalsheker, K Morgan

Alpha, antitrypsin or $\alpha_1$ proteinase inhibitor is the major serine proteinase inhibitor (serpin) present in the blood. An important physiological function of $\alpha_1$ antitrypsin is thought to be protection of the lower respiratory tract from damage by neutrophil elastase. Consequently, genetic deficiency of $\alpha_1$ antitrypsin is associated with a 20–50 fold increased risk of developing progressive lung damage, particularly in individuals who smoke. The mean plasma concentration is 1-4 g/l and the protein has a molecular weight of about 54 kilodaltons.

The protease-antiprotease theory of emphysema predicts that an imbalance between proteases, such as neutrophil elastase, and inhibitors, such as $\alpha_1$ antitrypsin, has an important role in the pathogenesis of chronic lung disease, and this hypothesis has been tested extensively in an attempt to obtain a better understanding of the mechanisms resulting in chronic lung damage.

This article will focus on (1) the application of molecular biology techniques to the study of the $\alpha_1$ antitrypsin gene; (2) the characterisation of molecular variants of $\alpha_1$ antitrypsin; and (3) the prospects for gene therapy.

Gene structure

The $\alpha_1$ antitrypsin gene is located on the long arm of chromosome 14 at position q31–31.2 Of interest is the fact that there are at least two homologous DNA sequences within about 200 kilobase pairs (kb) of the $\alpha_1$ antitrypsin gene, suggesting that they arose by duplication from an ancestral gene. One of the sequences codes for $\alpha_1$ antichymotrypsin and the other may be a non-functional gene or pseudogene, though there is a suggestion that the latter may also be expressed.

The $\alpha_1$ antitrypsin gene is present in 12.2 kb of genomic sequence and there are two promoter regions, one for monocytes and one for hepatocytes (fig 1). The two promoters are tissue specific, so that the monocyte promoter does not function in hepatocytes or the hepatocyte promoter in monocytes. This suggests that there are tissue factors or trans-activating factors that are required to activate the promoters in a tissue specific manner. In transgenic mice (containing the human $\alpha_1$ antitrypsin gene) $\alpha_1$ antitrypsin messenger RNA (mRNA) can be detected in several tissues, including the lung, Paneth cells in the gastrointestinal tract, and the kidneys, though the major source of plasma $\alpha_1$ antitrypsin is the liver. In the resting state human monocytes express about 1% of the amount of mRNA produced by hepatocytes, providing further support for the idea that the major source of plasma $\alpha_1$ antitrypsin in humans is also the liver.

The mRNA produced by monocytes differs from that produced by hepatocytes in that it contains additional exons and is present as two distinct mRNA species (fig 1). The smaller of the two transcripts results from the exclusion of the second monocyte exon (1B). Both transcripts apparently are present under basal conditions but the smaller form predominates after monocytes have been stimulated in culture by cytokines (unpublished observations). Probably the additional monocyte exons (1A and 1B) are transcribed but not translated into protein. The additional exons probably influence the rate at which protein is translated from the mRNA as the presence of untranslated mRNA tends to reduce the efficiency of translation. Although hepatocytes are the main source of plasma $\alpha_1$ antitrypsin, the relatively small amounts produced by monocytes may be an important local source of the protein in the immediate vicinity of the cells, particularly during inflammation. The first coding sequence for newly synthesised $\alpha_1$ antitrypsin protein is in exon II and it is likely, but not proved, that this sequence is identical in monocytes and hepatocytes.

Structure of protein

The mature protein secreted into the blood consists of 394 amino acids with three carbohydrate side chains attached at asparagine residues 46, 83, and 247. The tertiary structure of $\alpha_1$ antitrypsin has been extrapolated from studies of the mature protein after cleavage by enzymes. It is a globular, highly ordered molecule that consists of nine $\alpha$ helices (A–I) and three B pleated sheets. The active site of the protein is from amino acid residue 358 to 363 and two of these residues, methionine 358 and serine 359, appear to be critical in forming a “bait” for neutrophil elastase. From this information the likely mechanism of interaction between neutrophil elastase and $\alpha_1$ antitrypsin has been deduced (fig 2).
Deficiency states

Two common forms of \( \alpha_1 \) antitrypsin deficiency occur in populations of European origin—the S and Z variants (see the extensive reviews\(^{12-14} \)). Clinical manifestations (increased susceptibility to emphysema) appear to occur when the mean serum concentration of \( \alpha_1 \) antitrypsin is less than 35% of normal.\(^{13} \) About 1 in 3000 individuals in the United Kingdom are ZZ homozygotes and the frequency of the heterozygote carriers (MZ) is similar to that of individuals heterozygous for cystic fibrosis, which is widely regarded as the most common genetic disease in white people. Longitudinal studies in patients with the ZZ form of \( \alpha_1 \) antitrypsin deficiency indicate that the chance of being alive at the age of 50 is 52%, compared with 93% for the general population. By the age of 60 the chance of being alive is reduced to 16%, compared with 85% for the general population. In addition, for those with a history of cigarette smoking life expectancy is reduced by a further 10 years.\(^{13} \) These data indicate that inheritance of the ZZ genotype is associated with a reduced life expectancy, especially in smokers.

Since the original recognition of \( \alpha_1 \) antitrypsin deficiency about 75 molecular variants of \( \alpha_1 \) antitrypsin have been described, on the basis either of charge differences or of low serum concentrations. Many of these variants are rare and most have been characterised by the application of molecular biological techniques as arising from different DNA sequences.

In the earlier years of recombinant DNA technology some of the molecular variants were characterised by cloning and sequencing DNA, which thereby predicted the amino acid sequence. These procedures were relatively labour intensive and time consuming, though a vast improvement on previous methods, which relied on protein purification and amino acid sequencing. More recently, the advent of the polymerase chain reaction and sequencing of the amplified products has facilitated the characterisation of many more mutations relatively rapidly. The mutations of the \( \alpha_1 \) antitrypsin gene so far characterised are summarised in the table.

The ability to amplify specific sequences and analyse the products directly has also meant that specific deficiency states arising from single point mutations can be diagnosed rapidly, either by direct probing with oligonucleotides or by DNA sequencing, even when only small amounts of tissues are available for study.

Diagnosis of deficiency states

Short oligonucleotides (about 19 nucleotides in length) that contain even a single mismatched base pair in the middle are relatively unstable when hybridised to DNA sequences but highly stable when the match is perfect—that is, absolutely complementary. Consequently, DNA that contains a single point mutation can be identified if the conditions for hybridisation are optimised. In practice, duplicate filters are prepared and one is hybridised with a ‘wild type’ or normal probe and the other with the mutant probe of interest under optimised conditions. Stable hybridisation indicates a perfect nucleotide match. Effectively this means that the DNA does not have to be sequenced for mutations to be identified—they can be detected by using a labelled synthetic oligonucleotide, usually radioactive phosphorus (\( ^{32}P \))—that binds to a single stranded target DNA sequence on a solid membrane and exposing the membrane to autoradiography (fig 3).

Alternatively, a non-labelled oligonucleotide probe for the region of interest can be used as a primer in the reaction. Under optimised conditions gene amplification will occur only if the probe matches perfectly but not if there...
Normal and deficient variants of \( \alpha \)-antitrypsin

<table>
<thead>
<tr>
<th>Normal allele</th>
<th>Mutation</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1 (Ala(^{201}))</td>
<td>Ala(^{340})→Val</td>
<td>16</td>
</tr>
<tr>
<td>M1 (Val(^{201}))</td>
<td>Glu(^{340})→Asp, Arg(^{340})→His</td>
<td>17</td>
</tr>
<tr>
<td>M2</td>
<td>Glu(^{340})→Asp</td>
<td>18</td>
</tr>
<tr>
<td>M3</td>
<td>Arg(^{340})→His</td>
<td>19</td>
</tr>
<tr>
<td>M4</td>
<td>Arg(^{340})→Cys</td>
<td>20</td>
</tr>
<tr>
<td>Psaint albanis</td>
<td>Asp(^{340})→Asn</td>
<td>20</td>
</tr>
<tr>
<td>Vmunich</td>
<td>Asp(^{340})→Ala</td>
<td>21</td>
</tr>
<tr>
<td>Balhambra</td>
<td>Lys-Asp (location unknown)</td>
<td>22</td>
</tr>
<tr>
<td>X</td>
<td>Glu(^{340})→Lys</td>
<td>19</td>
</tr>
<tr>
<td>Xchristchurch</td>
<td>Glu(^{340})→Lys</td>
<td>23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Deficient allele</th>
<th>Mutation</th>
<th>Effect on secondary structure</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z</td>
<td>Glu(^{340})→Lys (Exon V)</td>
<td>Salt bridge lost</td>
<td>24</td>
</tr>
<tr>
<td>S</td>
<td>Glu(^{340})→Val (Exon III)</td>
<td>Salt bridge lost</td>
<td>25</td>
</tr>
<tr>
<td>Mheerlen</td>
<td>Pro(^{340})→Leu (Exon V)</td>
<td>Turning point instability</td>
<td>26</td>
</tr>
<tr>
<td>Mmalton</td>
<td>Phe(^{340})/Phe(^{340}) deleted (Exon II)</td>
<td>Folding abnormality</td>
<td>27</td>
</tr>
<tr>
<td>Mzgenotype</td>
<td>Gly(^{340})→Glut (Exon II)</td>
<td>Folding abnormality and functional deficiency</td>
<td>28</td>
</tr>
<tr>
<td>Mproicipida</td>
<td>Leu(^{340})→Pro (Exon II)</td>
<td>Helix distortion</td>
<td>29</td>
</tr>
<tr>
<td>Miiyama</td>
<td>Ser(^{340})→Phe (Exon II)</td>
<td>Turning point instability</td>
<td>30</td>
</tr>
<tr>
<td>I</td>
<td>Asp(^{340})→Cys (Exon II)</td>
<td>Salt bridge lost</td>
<td>27</td>
</tr>
<tr>
<td>Whedsheda</td>
<td>Ala(^{340})→Thr (Exon V)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ludwigshafen</td>
<td>Ile(^{340})→Asn (Exon II)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nulls</th>
<th>Mutation</th>
<th>Effect on secondary structure</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>granite falls</td>
<td>Tyr(^{340})→STOP (Exon II)</td>
<td>Truncated protein</td>
<td>31</td>
</tr>
<tr>
<td>bellingham</td>
<td>Lys(^{340})→STOP (Exon II)</td>
<td>Truncated protein</td>
<td>32</td>
</tr>
<tr>
<td>cardiff</td>
<td>Asp(^{340})→Val (Exon III)</td>
<td>Impaired folding</td>
<td>27*</td>
</tr>
<tr>
<td>newport</td>
<td>Glu(^{340})→Lys Gly(^{340})→Ser (Exon II)</td>
<td>Impaired folding +</td>
<td>33</td>
</tr>
</tbody>
</table>

| matarawa | Leu\(^{340}\) insert 3' shift→STOP (Exon V) | Truncated protein | 34 |
| isola de procida | Exons II→V deleted | Non-expression | 35 |
| hongkong | Leu\(^{340}\)/TC deleted 5' shift→STOP (Exon IV) | Truncated protein | 20 |
| bolton | Pro\(^{340}\) deleted, 5' shift→STOP (Exon V) | | |
| Zwexham | Glu\(^{340}\)→Lys, Ser\(^{340}\)→Leu (Exon I) Signal peptide | Impaired folding | 33 |

Mutation has also been identified as Pi Plowel\(^{36}\)

is a mismatch. Because an amplification of over a millionfold can be achieved sensitive radioactive probes are not required for detection. The major advantage in detecting \( \alpha \)-antitrypsin mutations by DNA technology is that any tissue, including blood from an individual or fetus, can be used as the polymerase chain reaction technique can amplify as little as a single gene copy.

Active site mutants
By using site directed mutagenesis \( \alpha \)-antitrypsin variants can be constructed to alter the coding sequences for the amino acids at the active site of the protein. This has resulted in some novel genetically engineered mutants with altered and potentially interesting biochemical properties. The two variants that have received most attention are methionine 358→valine and methionone 358→arginine. The former results in a form of \( \alpha \)-antitrypsin that still inhibits neutrophil elastase but is resistant to inactivation by oxidants such as those that are found in cigarette smoke or released by neutrophils. This property has resulted in the consideration of its role as a way of supplementing the lung’s ability to protect itself against damage by neutrophil elastase.\(^{36}\)
The second variant produces appreciably less elastase inhibition but is a major inhibitor of thrombin. Such a variant has a potential role in anticoagulation.

Clearly studies such as these clarify the relation between $z_1$ antitrypsin and its function and could result in the development of new therapeutic agents.

**Alpha, antitrypsin polymorphisms and respiratory disease**

It has been estimated that about 20% of patients with chronic airflow obstruction have a genetic component to their disease. As severe $z_1$ antitrypsin deficiency accounts for only about 2% of cases of chronic airflow obstruction, clearly other genetic factors contribute to the disease. Two polymorphisms of the $z_1$ antitrypsin gene have been identified in patients with chronic airflow obstruction by means of Southern blot and hybridisation analysis. These mutations occur in the untranslated 3' flanking region of the $z_1$ antitrypsin gene. There is some evidence that these two mutations account for an appreciable proportion of patients who develop chronic airflow obstruction. These polymorphisms may be associated directly with the expression of the $z_1$ antitrypsin gene or be associated with mutations in a neighbouring gene.

The TaqI polymorphism (fig 4) has been found in about 20% of unrelated patients with chronic airflow obstruction and occurs in the absence of plasma $z_1$ antitrypsin deficiency. This polymorphism conferred an estimated 13 fold risk for developing chronic airflow ob-

---

**Gene therapy**

The Z form of $z_1$ antitrypsin arises from a single point mutation in the gene for the normal allele. This alters the codon corresponding to position 342 in the protein, so that glutamic acid is replaced by lysine. Predictions made from the tertiary structure of cleaved $z_1$ antitrypsin suggest that the mutation in the Z form results in abnormal folding and accumulation of the protein in the endoplasmic reticulum. A salt bridge normally exists between amino acid residues 290 (lysine) and 342 (glutamic acid). This salt bridge is thought to be lost in the Z variant when the amino acid 342 becomes lysine, and may cause the failure to secrete the protein.

There are conflicting data on whether the loss of the salt bridge adequately explains defective secretion. Attempts have been made to correct the Z defect by site directed mutagenesis. A second mutation was introduced into a complementary DNA (cDNA) clone corresponding to the Z gene to recreate a salt bridge between amino acids 342 and 290 in the protein. The codon for residue 290 has been converted from lysine to glutamic acid to reverse the two residues found in the normal gene, thereby replacing the salt bridge with the Z variant lysine at 342.

Transfection of mammalian cells with such constructs have yielded conflicting results. One study suggested that the secretory defect could be reversed whereas two other studies have suggested that no correction of the defect occurs. The Z variant can also be converted back to normal, however, by replacing the abnormal lysine at 342 with the normal glutamic acid. The resulting product is processed and secreted normally, indicating the critical nature of amino acid 342. Correction of abnormal genes is potentially possible, though other approaches to gene therapy may prove more practical.

Attempts have been made to investigate retroviral mediated transfer of the normal $z_1$ antitrypsin gene in genetic mammalian cells in culture. Although the protein can be integrated and expressed in such cells, this approach is also unlikely to be practised at present. There are several potential problems, including inability to regulate expression of the gene, and integration into the cellular genome is likely to be random and may disrupt or activate other genes.

**Replacement therapy with purified $z_1$ antitrypsin**

Clinical trials in patients with $z_1$ antitrypsin...
deficiency are currently under way. Replacement therapy with α1 antitrypsin purified from plasma has been used, and it has been shown to generate adequate plasma and lung concentrations when given parenterally or by inhalation.14-16 Because of the slowly progressive nature of emphysema, even in patients with α1 antitrypsin deficiency, a long period and many patients will be needed for the benefits of such treatment to be assessed. In early studies, however, no appreciable side effects occurred and biochemical efficacy was observed, though treatment has to be given monthly parenterally or twice daily by inhalation. Thus it has been suggested that subjects should be selected on the basis of having unequivocal evidence of lung disease with an FEV1 of under 65% of the predicted normal value, should have stopped smoking and should have been followed up for at least two years for documentation of a substantial decline in FEV1 (over 80 ml a year).17 Because of the large number of patients with α1 antitrypsin deficiency, it will prove infeasible to quantify enough protein from plasma to fulfil the potential need. For this reason recombinant α1 antitrypsin produced commercially may be required. Alternatively, treatment may be unnecessary if patients have stopped smoking and may be required only for episodes of acute infection when neutrophils and their elastase load are increased in the lung. Other potential applications of α1 antitrypsin, and indeed other synthetic elastase and protease inhibitors, include the treatment of conditions such as septic shock and adult respiratory distress syndrome, where protease release is increased and implicated in pathogenesis.

Our understanding of the molecular basis of respiratory disease has been influenced substantially by recombinant DNA technology. The possibility of characterising mutations and detecting them early and the potential for providing treatment illustrate how the application of such techniques could have an important impact on the morbidity associated with certain types of chronic respiratory disease and should also provide new insights into their pathogenesis.

Molecular biology and respiratory disease. 7. The alpha 1 antitrypsin gene and chronic lung disease.
N Kalsheker and K Morgan

Thorax 1990 45: 759-764
doi: 10.1136/thx.45.10.759

Updated information and services can be found at:
http://thorax.bmj.com/content/45/10/759.citation

These include:

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/